Índice general

1. Introducción **3**
- 1.1. Algunos Ejemplos . . . 3
- 1.2. El entorno matematico . . . 6
- 1.3. La variedad de los problemas de optimización . . . 12
- 1.4. Ejercicios . . . 13

2. Programación Lineal **19**
- 2.1. Introducción . . . 19
- 2.2. El metodo simplex . . . 24
- 2.3. Dualidad . . . 35
- 2.4. Algunos asuntos practicos . . . 39
- 2.5. Programación entera . . . 47
- 2.6. Ejercicios . . . 50

3. Programación no lineal **55**
- 3.1. Problema modelo . . . 55
- 3.2. Multiplicadores de Lagrange . . . 57
- 3.3. Condiciones de optimalidad Karush-Kuhn-Tucker . . . 64
- 3.4. Convexidad . . . 70
- 3.5. Dualidad y convexidad . . . 79
- 3.6. Ejercicios . . . 83

4. Tecnicas de Aproximación **87**
- 4.1. Introducción . . . 87
- 4.2. Métodos de busqueda lineal . . . 89
- 4.3. Metodos de gradiente . . . 91
- 4.4. Metodos de Gradiente Conjugado . . . 94
- 4.5. Aproximación bajo restricciones . . . 98
- 4.6. Comentarios Finales . . . 104
- 4.7. Ejercicios . . . 105

5. Problemas Varacionales y Programación Dinamica **109**
- 5.1. Introduccion . . . 109
- 5.2. La ecuación de Euler-Lagrange: Ejemplos . . . 111

5.3. La ecuación de Euler-Lagrange: Justificación 121
5.4. Condiciones Naturales de Frontera 124
5.5. Problemas variacionales bajo condiciones integrales y puntuales . 126
5.6. Resumen de las restricciones para problemas variacionales 132
5.7. Problemas variacionales de diferente orden 135
5.8. Programación Dinamica: La ecuación de Bellman 139
5.9. Ideas basicas en la aproximación númerica 143
5.10. Ejercicios 146

6. Control Óptimo **151**
6.1. Introducción 151
6.2. Multiplicadores y el Hamiltoniano 152
6.3. El principio de Pontryagin 158
6.4. Otro Formato 173
6.5. Algunos Comentarios en la Aproximación Numérica 174
6.6. Ejercicios 178

Capítulo 1

Introducción

1.1. Algunos Ejemplos

Creemos que no hay mejor manera de convencer al lector del interes y la aplicabilidad de ciertas ideas o tecnicas matematicas que mostrar el tipo de problemas practicos que pueden ser abordados y eventualmente resueltos usando estas ideas. Al mismo tiempo, esta lista inicial de ejemplos y problemas es una clara muestra de los objetivos del texto. Algunos de los ejemplos pueden no ser entendibles en una primera lectura, lo cual no debe preocupar a nuestros lectores ya que insistitemos en ellos a traves de este capitulo y su significado estara mas claro al final del mismo. La mayoria de los ejemplos que estudiaremos son muy conocidos y academicos, en el sentido que el tamao de los problemas reales no es comparable con las situaciones que estudiaremos. Versiones mas completas de estos problemas pueden ser encontradas en textos mas avanzados. Sin enbargo, consideramos que las ideas planteadas con estos daran al lector las herramientas basicas para situaciones mas realistas.

El problema del transporte *Se va a enviar cierto producto en cantidades* $u_1, u_2, \ldots, u_n$ *desde* n *estaciones de servicio a* m *destinatarios, donde se deben recibir en cantidades* $v_1, v_2, \ldots, v_m$ *Vease figura 1.1. Si el costo de enviar una unidad de producto desde el origen* i *al destino* j *es* c_{ij}*, determine la cantidad* x_{ij} *a enviar del origen* i *al destino* j *de tal manera que el costo total de transporte es minimo.*

El problema de la dieta *Se conocen los contenidos nutricionales asi como los precios de ciertos alimentos. Asi tambien la cantidad minima requerida de cada nutriente. La tarea consiste en determinar la cantidad de cada alimento que se debe adquirir de tal manera que se obtenga la cantidad minima de cada nutriente y que el costo total de la dieta sea el menor posible.*

El sistema de andamiaje *Considere el sistema de andiamaje de la figura 1.2, donde las cargas* x_1 *y* x_2 *son aplicadas a ciertos puntos de las vigas* 2 *y*

3 *respectivamente. Las cuerdas A y B pueden soportar como maximo* 300 *kg cada una, las cuerdas C y D pueden soportar* 200 *kg cada una, y las cuerdas E y F, un maximo de* 100 *kg cada una. Encontrar la carga maxima* $x_1 + x_2$ *que puede soportar el sistema manteniendo equilibrio de fuerzas y equilibrio de momentos, las cargas optimas* x_1 *y* x_2*, y los puntos optimos donde deben ser aplicadas. Asumiendo que el peso de las cuerdas y de las vigas es despreciable.*

Estimacion de la potencia en un circuito *Las variables de estado en una red electrica son los voltajes en cada nodo de la red, cada uno un numero complejo con modulo* v_i *y argumento* δ_i*. Las potencias activas y reactivas en la coneccion entre los nodos* i *y* j *esta dada por:*

$$p_{ij} = \frac{v_i^2}{z_{ij}} \cos\theta_{ij} - \frac{v_i v_j}{z_{ij}} \cos(\theta_{ij} + \delta_i + \delta_j)$$
$$p_{ij} = \frac{v_i^2}{z_{ij}} \sin\theta_{ij} - \frac{v_i v_j}{z_{ij}} \sin(\theta_{ij} + \delta_i + \delta_j)$$

donde el modulo z_{ij} *y la fase* θ_{ij} *determinan la impedancia de la linea* ij*. Si poseemos mediciones experimentales* $\bar{v_i}$,$\bar{p_{ij}}$,$\bar{q_{ij}}$ *de los respectivos valores* v_i,p_{ij} *y* q_{ij}*, y los parametros del error en la medición son* k_i^v*,* k_{ij}^p *y* k_{ij}^q *respectivamente, estimar el estado de la red minimizando, en las variables* v_i *el error medio cuadratico de las mediciones disponibles con respecto a los valores predichos de tal manera que las anteriores formulas se cumplan de la mejor forma posible.*

Diseño de un solido en movimeinto *Se desea diseñar un solido con simetria radial alrededor de cierto eje que debe viajar en linea recta con velocidad constante dentro de un fluido. Si la densidad del fluido es lo suficientemente pequea, entonces el modulo de la presion normal en la dirección de la normal exterior a la superficie del cuerpo ejrcida por el fluido sobre el solido esta dada por:*

$$p = 2\rho v^2 \sin^2\theta$$

donde ρ *y* v *son la densidad y la valocidad del fluido relativo al solido (ambas constantes), y* θ *es el angulo formado por el tangente de el perfil de la superficie en el plano* xy *y la velocidad del fluido (ver figura 1.3). ¿Comó podemos encontrar el perfil optimo del solido para minimizar la presión que ejerce el fluido en este?*

Diseño de un canal *Los canales son un tipo particualar de dispositivo para transportar fluidos. Normalmente el fluido no ocupan todo el canal (Figura 1.4) y en general las perdidas se originan en las paredes. En un modelo especifico la fricción puede ser aproximada por:*

$$\frac{1}{\sqrt{f}} = 2\log\frac{3{,}7D_h}{e}$$

donde f *es el coeficiente de fricción.* D_h *es el llamado diametro hidraulico, y* e

representa una medida de rugosidad. Mas aun, tenemos:

$$D_h = 4R_h, \qquad R_h = \frac{A}{P}$$

Donde A es el area de la sección transversal del canal ocupada por el fluido, y P es el perimetro alcanzado por la misma sección transversal de fluido. Si asumimos que A es fijo, el objetivo del problema es determinar la forma de la sección transversal del canal que minimizara las perdidas de fluido a traves de las paredes.

El constructor de botes *Un constructor de botes tiene los siguientes comprmisos durante el ao: Al final de Marzo un bote, en Abril dos, en Mayo 5, al final de Junio 3, durante Julio 2 y en Agosto 1. El puede contruir como maximo cuatro botes por mes, y puede tener en "stock" como maximo tres. El costo de cada bote es 10.000 euros, mientras que mantener uno en "stock" cuesta 1.000 euros mensuales. ¿Cuál es la estrategia optima para contruir los botes y minimizar los costos?*

Oscilador armonico con frición *El control de superficie en un objeto volador debe ser mantenida en equilibrio en cierta posición. Las fluctuciónes mueven la superficie, y si no son contrarestadas, esta vibrará cumpliendo la siguiente ecuación:*

$$\theta'' + a\theta' + w^2\theta = 0$$

donde θ es el angulo medido desde la posición de equilibrio deseado, y a y w son constantes dadas. Un servomecanismo aplica un toque que cambia el comportamiento de la oscilación a:

$$\theta'' + a\theta' + w^2\theta = u$$

donde el control u debe estar acotado (i.e $|u(t)| \leq C$). El problema consiste en determinar el parametro del servomecanismo $u(t)$ de tal manera que la superficie regrese a reposo $\theta = \theta' = 0$ desde un estado arbitrario $\theta = \theta_0$, $\theta' = \theta'_0$ en tiempo minimo.

El problema de la posición *Cierto objeto movil en el plano es controlado por dos parametros: La magnitud de la aceleración γ y la razon de cambio del angulo de rotación θ'. Si asumimos que γ y θ' se les permite moverse en los inervalos $[-a, a], [-\alpha, \alpha]$ respectivamente, determinar la estrategia optima para llevar el objeto movil desde unas condiciones inicales, al reposo en el origen.*

A pesar que la colección de problemas y ejemplos puede ser ampliada considerablemente (incluyendo ejemplos mas cercanos a la realidad o a las situaciones tecnologicas o de ingenieria), los anteriormente mencionados nos dan la idea que estamos frente a un tema de caracter aplicado. Aprenderemos a enfrentar y resolver estos problemas y muchos otros en los siguientes capitulos. Una vez se hallan entendido estas ideas, el lector sera capaz de analizar y resolver por

si mismo muchas otras situaciones que se originan en la ciencia y la tecnologia. Tambien podria profundizar su conocimiento de una clase particular de problemas consultando textos mas avanzados en esa area.

1.2. El entorno matematico

Los ejemplos de la sección anterior son aparentemente muy diferentes entre si, pero sin embargo todos comparten algo en comun que les permite que se les presente en este libro. En todas estas situaciones estamos buscando una solucion optima, la mejor manera de hacer las cosas, la mas eficiente, el proceso mas economico. Debido a esto, todas las ideas que se han desarrollado durante los años para examinar estos problemas pueden ser puestas bajo la categoria de OPTIMIZACIÓN. Sin embargo los problemas anteriores son muy diferentes los unos de los otros, y las tecnicas para resolverlos o aproximar sus soluciones reflejan esta misma variedad. No se espera que en este momento los lectores descubran estas diferencias por si mismos, estas diferencias, incluso mas antes de colocarlas de una forma mas precisa, cuantitativa reflejan fielmente cada situacion y permiten un tratamiento adecuado llevando a la solución o a una buena aproximación numerica de esta. Este proceso de ir desde la formulación en palabras de una situacion particular a su formulacion en tarminos matematicos precisos es de tal importancia que la imcapacidad de llevarlo a cabo correctamente puede llevar a respuestas absurdas a los problemas. El exito al usar cierta tecnica de optimizacion depende claramente de este paso.

La formulación del problema en terminos matematicos precisos debe reflejar exactamente lo que se desea resolver. En particular, cuando se trabaja con problemas de optimización hay dos pasos importantes a seguir. Primero, la función objetivo o costo debe reflejar de forma veraz nuestra idea de optimo. Una solucion deseable debe tener el menor (o mayor) costo funcional, ser un tiempo minimo, una mayor eficiencia, mayores beneficios, perdidad minimas, etc. Si nuestro funcional de costo no refleja correctamente nuestro criterio de optimización, la solución final no sera presumiblemente la solución optima buscada. Segundo, es igualmente importante establecer de forma explicita las restricciones que deben ser impuestas en el problema, de tal forma que las soluciones admitibles son verdaderamente posibles en nuestro problema o situación. De nuevo, si estas restricciones no son escritas de forma adecuada, algunas de ellas son olvidadas o estamos imponiendo demasiadas de manera que somos demasiado restrictivos, nuestra respuesta final puede no ser lo que estamos buscando. Con el objetivo de enfatizar estos puntos vamos a tratar sucesivamente, los problemas anteriores y dar su formulación matematica. Antes de proceder a la tarea, mencionemos algunas cosas que hay que tener presentes cuando se enfrenta alguna situacion particular.

Hemos enfatizado la importancia de la trancisión de un problema de optimización, a menudo dado en plabras, a su formulación cuantitativa precisa que nos permitira eventualmente resolver el problema. Los cientificos y los ingerieros deberian volverse expertos en este proceso. Algo que no se debe olvidar cuando

se esta planteando o replanteando una situación es insistir que en cada estado de la formulación se vea reflejado nuestro objetivo original de tal manera que la conección entre la situación a ser resuelta y su formulación este siempre allí. Esto requiere una actitud activa con respecto a la formulación o reformulación de un problema particular hasta que hallamos interpretado cada aspecto de la situación.

Para evitar que estos comentarios sean inutiles nos atrevemos a dar las siguientes recomentaciones a aquellos enfrentandose a un problema de optimización.

Entendiendo el criterio de optimización *Debe haber un entendimiento claro del objetivo y de la manera la cual la optimalidad va a ser medida. En particular, la desición de cuales variables depende el costo y las restricciónes sobre las mismas es crucial. Un problema puede ser planteado de varias maneras distintas, y es importante distinguir cual de estas es la forma mas eficiente de plantearlo. Es importante revisar los valores extremos de las variables (y otros valores relevantes) y si el costo asiciado es coherente con lo esperad. Este tipo de analisis usualmente lleva al descubrimiento de errores acerca del problema, y a una revisión de las variables, restricciones y funcional objetivo.*

Entendiendo las restricciones *Las restricciones que vinculan las variables de distintas maneras son igualmente significativas. Estas pueden ser de naturaleza muy diferente: igualdades, desigualdades, equaciones diferenciales, restricciones integrales, etc. Estas pueden estar ocultas de muchas maneras, en ocaciones de forma explicita o implicita. Lo que es importante es analizar las relaciones entre las variables y las restricciones que se deben respetar. En particular las igualdades pueden ser utilizadas para disminuir el numero de variables. La misma actitud descrita anteriormente nos debe llevar a revisar las restricciones y su coherencia con respecto a la situacion que deseamos examinar.*

Reflejando la formulación precisa *Una vez se hallan cubierto los dos pasos anteriores, vale la pena revisar la formulación matematica del problema. ¿Son las resticciones coherentes? ¿Podría el conjunto de soluciones aceptables estar vacio? ¿Podrian algunas de las restricciones ser simplificadas o eliminadas porque algunas de las restricciones son mas fuertes que otras? ¿Puede el costo hacerse tan pequeño como se quiera sin violar las resticciones? Si es asi es posible que hallamos olvidado alguna restricción. ¿Podemos anticipar si hay una unica solución optima o si hay varias?*

Analisis rapido de la solución *Finalmente, es buena idea acostumbrarse a examinar brevemente la solución optima que se ha obtenido o aproximado. ¿Parece que ofrece el minimo costo, maxima eficiencia, etc.? ¿Es posible que esta en realidad es una solucion optima? ¿Esta refleja la optimalidad con respecto a las condiciones del problema inicial? ¿Esta cumple con todos los requerimientos?*

Como dice el dicho "La practica hace al maestro", y los problemas y tecnicas de optimización no son la exepción. Los ejercicios y problemas ayudaran a los

estudiantes a ir a traves de todos los pasos descritos anteriormente de forma radida y acertada. Al principio habra inseguridad, errores, ineficiencia, falta de ideas para superar dificultades etc, pero a medida que los estudiantes dominan estos aspectos iran creando auto-confianza.

A continuación damos la formulación precisa de los diferentes problemas propuestos en la sección anterior. Les pedimos a los estudiantes que trabajen en entender la conección entre la formulación original del problema y su traducción a ecuaciones, formulas, desigualdades, igualdades etc. Este proceso normalmente incluye desarrollar el modelo para la situación dada. En algunos casos simples, el modelo sera lo suficientemente claro, y no surgira ninguna dificultad para poner el problema en una forma apropiada. En otros sin embargo puede haber un a dificultad inicial para entender los mecanismos asociados con una situación especifica, y se requerira un esfuerzo adicional para comprender su siginificado y alcanzar una formulación precisa.

El problema del transporte *Si x_{ij} es la cantidad de producto enviado desde la posición inicial i al destino j, el costo total sera:*

$$\sum_{i,j} c_{ij} x_{ij}$$

si c_{ij} es el costo de enviar una unidad de producto de i a j. '?Cualés son las restricciones que tenemos que respetar? Para un punto de servicio fijo i, u_i es la cantidad a ser enviada, asi que

$$\sum_{j} x_{ij} = u_i, \qquad i = 1, 2, \ldots, n$$

de igual manera, para cada destino fijo se debe recibir la cantidad v_j, esto nos obliga a

$$\sum_{i} x_{ij} = v_j, \qquad j = 1, 2, \ldots, n$$

Hay que notar que estos dos conjuntos de igualdades son compatibles si

$$\sum_{i} u_i = \sum_{j} v_j$$

que es una restricción que los datos del problema deben cumplir para que el problema tenga sentido. Ademas, si damos por dado que el hecho de ser un punto de servicio o un destino no puede er revertido debemos pedir que

$$x_{ij} \geq 0, \qquad \textit{para todo } i, j$$

Ahora lo que estamos buscando es:

$$\textit{Minimizar} \qquad \sum_{i,j} c_{ij} xij$$

bajo

$$\sum_j x_{ij} = u_i, \qquad i = 1, 2, \ldots, n$$
$$\sum_i x_{ij} = u_j, \qquad j = 1, 2, \ldots, m$$
$$x_{ij} \geq 0 \qquad \text{para todo } i, j$$

El problema de la dieta *Sea x_i la cantidad de comida i que se va a comprar. El costo total que que quisieramos es*

$$\sum_i c_i x_i$$

si c_i es el precio de una unidad de comida i. Sea a_{ij} el contenido del nutriente j por unidad de comida i, y b_j la cantidad minima requerida del nutriente j. Entonces debemos asegurarnos que en la dieta que escojamos cumple este minimo:

$$\sum_i a_{ij} x_i \geq b_j, \qquad \text{para todo } j$$

Finalmente debemos pedir que cada x_i es no negativo:

$$x_i \geq 0, \qquad \text{para todo } i$$

El problema es:

$$\text{Minimizar} \qquad \sum_i c_i x_i$$

sujeto a:

$$\sum_i a_{ij} x_i \geq b_j, \qquad \text{para todo } j$$
$$x_i \geq 0, \qquad \text{para todo } i$$

El sistema de andamiaje *Si notamos por T_A, T_B, T_C, T_D, T_E y T_F las tensiones en cada cuerda cuando llevan cargas x_1 y x_2 aplicadas en los puntos ubicados a x_3 y x_4 unidades de distancia de el borde izquierdo de la respectiva viga, las condiciones de equilibrio de fuezas y momentos nos llevan a las equaciones*

$$T_E + T_F = x_2, \qquad 8T_F = x_4 x_2,$$
$$T_C + T_D = x_1 + T_E + T_F, \qquad 10T_D = x_3 x_1 + 2T_E + 10T_F,$$
$$T_A + T_B = T_C + T_D, \qquad 12T_B + 2T_C + 12T_D.$$

Si ahora expresamos las diferentes tensiones en cada cuerda con respecto a nuestras variables de diseño x_i tenemos

$$\frac{x_2x_4}{8} = T_F \leq 100, \qquad \frac{8x_2 - x_2x_4}{8} = T_E \leq 100,$$
$$\frac{2x_2 + x_1x_3 + x_2x_4}{10} = T_D \leq 200, \qquad \frac{10x_1 + 8x_2 - x_1x_3 - x_2x_4}{10} = T_C \leq 200,$$
$$\frac{2x_1 + 4x_2 + x_1x_3 + x_2x_4}{12} = T_B \leq 300, \qquad \frac{10x_1 + 8x_2 - x_1x_3 - x_2x_4}{12} = T_A \leq 300.$$

por lo tanto se deben complir estas desigualdades. Ademas debemos pedir que

$$x_1 \geq 0, \qquad x_2 \geq 0, \qquad 0 \leq x_3 \leq 10, \qquad 0 \leq x_4 \leq 8$$

El problema es entonces.

$$\textit{Maximizar} \qquad x_1 + x_2$$

sujeto a:

$$x_1 \geq 0, \qquad x_2 \geq 0,$$
$$0 \leq x_3 \leq 10, \qquad 0 \leq x_4 \leq 8,$$
$$x_2x_4 \leq 800, \qquad 8x_2 - x_2x_4 \leq 800,$$
$$2x_2 + x_1x_3 + x_2x_4 \leq 200, \qquad 10x_1 + 8x_2 - x_1x_3 - x_2x_4 \leq 3600.$$

Estimacion de la potencia en un circuito *En este ejemplo se nos pide minimizar el error medio cuadratico de ciertas medidas con respecto a los valores predichos. Especificamente buscamos*

$$\textit{Minimizar} \qquad \sum_{i\in\Omega} k_i^v (v_i - \bar{v_i})^2 + \sum_{i\in\Omega}\sum_{j\in\Omega_i} k_{ij}^p (p_{ij} - \bar{p_{ij}})^2 + \sum_{i\in\Omega}\sum_{j\in\Omega_i} k_{ij}^q (q_{ij} - \bar{q_{ij}})^2$$

donde los distintos datos estan dados en el enunciado y

$$p_{ij} = \frac{v_i^2}{z_{ij}}\cos\theta_{ij} - \frac{v_iv_j}{z_{ij}}\cos(\theta_{ij} + \delta_i + \delta_j)$$
$$p_{ij} = \frac{v_i^2}{z_{ij}}\sin\theta_{ij} - \frac{v_iv_j}{z_{ij}}\sin(\theta_{ij} + \delta_i + \delta_j)$$

Las variables desconocidas son (v_i, δ_i) y no tenemos una restricción especifica en esto. Aqui Ω es el conjunto de nodos, Mientras que Ω_i es el conjunto de aquellos conectados al nodo i.

Diseño de un solido en movimiento *De acuerdo a nuestra explicaión anterior y al digrama correspondiente, el componente paralelo al eje x de la presión normal en un punto de la superficie del solido es*

$$\rho\sin\theta = 2\rho v^2\sin^3\theta$$

La presión total en una rebanada de ancho dx sera el producto de la expresión anterior con la superficie lateral de la rebanada.

$$dP = 2\rho v^2 \sin^3\theta 2\pi y(x)\sqrt{1+y'(x)^2}dx$$

si un perfil dado es obtenido rotando la grafica de la función $y(x)$, y escribimos $\sin\theta$ en terminos de $\tan\theta = y'(x)$ obtenemos

$$dP = 2\rho v^2 2\pi \frac{y'(x)^3}{(1+y'(x)^2)^{\frac{3}{2}}} y(x)\sqrt{1+y'(x)^2}dx$$

o simplificando

$$dP = 4\pi\rho v^2 \frac{y(x)y'(x)^3}{1+y'(x)^2}dx$$

y estamos interesados en encontrar el perfil $y(x)$ que minimice la integral anterior entre todas las funciones (continuas) que satisfagan $y(0) = 0$, $y(L) = R$.

Diseño de un canal *Dado que las perdidas en las paredes del canal son proporcionales al inverso del perimetro, para un area de sección dada fija, el mejor perfil es el que tiene el menor perimetro posible. Mas especificamente estamos buscando el perfil $y(x)$ que minimiza la integral.*

$$\int_0^R \sqrt{1+y'(x)^2}dx$$

que da la longitud de la grafica de $y(x)$, sujeto a

$$y(0) = 0, \qquad y(R) = 0, \qquad \int_0^R y(x)dx = A$$

El constructor de botes *Este problema es suficientemente claro y no es necesario una explicacion mas clara.*

El oscilador armonico con fricción *En este ejemplo, el mejor control $u(t)$ es aquel que lleva la superficie oscilatoria al reposo lo mas pronto posible y al mismo tiempo cumple la restricción en el tamaño $|u(t)| \leq C$*

El problema de la posición *Un objeto movil en el plano puede ser controlado por dos parametros a nuestra disposición, r_1 y r_2, expresando el modulo de cambio de velocidad y rapidez con la cual la dirección del movimiento se puede cambiar (velocidad angular del movimiento) respectivamente. Las ecuaciones de movimiento son*

$$x''(t) = r_1(t)\cos(\theta(t)), \qquad y''(t) = r_1(t)\sin(\theta(t)), \qquad \theta'(t) = r_2(t)$$

Las restricciones en los pares aceptables (r_1, r_2) se escriben requiriendo que

$$(r_1, r_2) \in [-a, a]x[-\alpha, \alpha]$$

El objetivo es cambiar la posición del objeto en reposo ($x'(0) = y'(0) = 0$) desde un punto arbitrario (x_0, y_0). Al origen en un tiempo minimo

$$x(T) = y(T) = x'(T) = y'(T) = 0$$

done T es lo mas pequenõ posible.

1.3. La variedad de los problemas de optimización

Hemos hecho notar, y es probable que el lector haya notado, las inmensas diferencias entre los distintos tipos de problemas de optimización. Estas diferencias han motivado la estructura de este texto.

Posiblemente la diferencia mas significativa esta en el hacho que en algunos problemas son vectores los que describen soluciónes y soluciones optimas. Mientras que en otros casos son funciones lo que se necesita para plantear y resolver el problema. Esta distincion importante profunda y cualitativa da lugar a diferentes tecnicas de optimizacion usadas en estas dos categorias de problemas. La situación es similar al caso de ecuaciones o sistemas de ecuaciones en los cuales buscamos un vector como solución, o sea simplemente numeros, y a ecuaciones diferenciales donde la incognita es una función. En el primer caso hablamos de programación matematica; en el segundo acerca de problemas variacionales. Siguiendo con esta clasificación la programación matematica se puede dividir en programación lineal (Capitulo 2), que trabaja en el mundo (mas simple) de los problemas lineales, y programación no lineal (Capitulo 3) para las complejas tecnicas de optimización no lineal. Los problemas del transporte y de la dieta corresponden a programación lineal, mientras que los problemas del andamiaje y la estimación del circuito corresponden a la programación no lineal.

Las situaciones donde intentamos hallar funciones optimas para una situacion dada pueden ser clasificadas en problemas variacionales (Capitulo 5), programación dinamica y control optimo (Capitulo 6). Los problemas de diseño de un solido en movimiento, el de diseño del canal y del constructor de botes corresponden a problemas variacionales y programación dinamica. El ,oscilador armonico y el problema de la posición son problemas tipicos de control optimo.

El capitulo cuarto es un punto de intersección entre el mundo de los vectores y el mundo de las funciones. Esto sera claro mas tarde. Nuestro objetivo en este capitulo es describir los algoritmos numericos mas basicos y relevantes para calcular y/o aproximar soluciones a los problemas. Dado que en la mayoria de las situaciones del mundo real no se esperan soluciones exactas, estas herramientas computacionales son cruciales. Restringiremos nuestra atencion a las tecnicas mas basicas y bien conocidas. Nuestra idea es que el lector tenga alguna idea de la naturaleza de las tecnicas de aproximación en los problemas de optimización. No se ha incluido la implementación explicita de los algoritmos por dos motivos. El primero la cantidad de software comercial comprobado (Ver el capitulo 4 para referencias), que son bastante utiles dado que nos liberan de los detalles tecnicos relacionados con la aproximación, y nos permiten concentrarnos en el

modelamiento. Por otro lado, la calibración fina de los algoritmos, espacialmente cuando se consideran restricciones no lineales, requiere considerable experiencia apenas el numero de variables crece mas alla de unas cuantas. El inexperto haria un trabajo pobre comparado con el que haria uno de estos paquetes de software. Esto no quiere decir que es inutil adquirir algo de experiencia tratando de escribir programas para algunas situaciones simples. Por lo cual hemos escrito algunas versiones simples de los algoritmos en el formato de seudocodigo.

Finalmente es importante resaltar que cada uno de estos capitulos no es mas que una timida introducción a las ideas correspondientes. La diversidad de situaciones, las peculiaridades de los problemas reales, la necesidad de mejores algoritmos y metodos computacionales, y la necesidad de una comprensión mas profunda de la naturaleza de los problemas puede ser tal que un libro completo seria necesario para cubrir cada uno de estos pequeños capitulos. Nuestra intención es dar una primera vista general a la optimización, emfatizando las ideas basicas y las tecnicas en cada categoria de problemas de optimización.

1.4. Ejercicios

1. Un inversionista busca invertir cierta cantidad de capital K de una manera diversificada para maximizar la ganancia esperada despues de cierto periodo de tiempo. Si r_i es la taza de interes promedio para la inversion i, y para evitar riesgo execivo, no se quiere en cualquier inverción mas de un porcentaje fijo r del capital. Formule el problema que lleve a la mejor solucion. ¿Puede pensar en otras restricciones razonables para esta situación

2. En el contexto del problema del andamiaje descrito anteriormente en este capitulo, asuma que los puntos donde las cargas x_1 y x_2 son aplicadas son exactamente los puntos medios de las vigas CD y EF respectivamente. Formule el problema. ¿Cual es la diferencia principal entre esta situación y aquella descrita en el texto?

3. Una compañia fabricante de tejas debe proveer 7800 m^2 de estas para varias casas. Dos diferentes tipos de tejas pueden ser usados. El modelo A10 requiere 9.5 tejas por metro cuadrado el modelo A13 necesita 12.5 elementos por metro cuadrado. Ambos modelos pueden ser usados en el mismo techo. Los respectivos precios son 0.70 y 0.80 euros por elemento. La compañia tiene 1600 horas laborales para terminar los techos. En una hora 5 m^2 del modelo A10 y 4 m^2 del modelo A13 pueden ser instalados. Debido a restricciones de inventario la maxima cantidad del modelo A13 es 2500 m^2. Formule el problema de maximizar los beneficios sujeto a todas las restricciones indicadas.

4. En el sistema de resortes de la figura 1.5, cada nodo es libre de rotar sobre si mismo. La constante de elasticidad de cada resorte es k_i (de acuerdo a la ley de Hook), y la posición de equilibrio del nodo central es determinada

mediante el sistema

$$\sum_i k_i(x - x_i) = 0$$

donde x_i es la posición de los nodos fijos, desciba como determinar las constantes optimas de los resortes k_i, que minimizan el trabajo hecho por una fuerza constante F en el nodo libre, asumiendo que

$$\sum_i k_i = k$$

es una constante positiva fija.

5. Una compañia se dispone a construir algunos (m) puntos de servicio para atender a cierto numero (n) de clientes conocidos. Se debe decidir la posición de estos puntos de servicio. Asumiento que el criterio escogido es minimizar la distancia global de los puntos de servicio a los clientes, formule el problema como uno de programación no lineal. Describa otras maneras de tomar ese desición.

6. Una formula de cuadratura es una manera de aproximar eficientemente integrales definidas a traves de sumas del tipo

$$\int_a^b f(x)dx \approx \sum j = 1^n \alpha_j f(x_j)$$

Donde los pesos α_j y los puntos x_j determinan una forma particular de cuadratura. Queremos determinar el vector de n pesos (α_j) y n puntos (x_j) en el intervalo $[-1, 1]$ de tal manera que la cuadratura es exacta para polinomios de grado lo mas alto posible. El procedimiento consiste en minimizar el error cuadratico de la formula de cuadratura para polinomios de grado m. Plantee el problema como un problema de optimización no lineal.

7. La función de utilidad de Cobb-Douglas es de la forma

$$u(x, y) = x^\alpha y^{1-\alpha}, \qquad 0 < \alpha < 1, \qquad x \geq 0, \qquad y \geq 0$$

Asuma una economia de dos consumidores 1 y 2, y dos bienes X y Y. Ambos consumidores tienen la misma función de utilidad del tipo mencionado anteriormente con el mismo exponente α, y recursos

$$(\bar{x}_i, \bar{y}_i), \qquad i = 1, 2$$

para cada bien. Si los precios $p = (p_x, p_y)$ prevalecen en el mercado para cada bien, formule el problema de maximizar la satisfacción de cada consumidor medida por sus funciones de utilidad.

8. Se debe recargar una escalera contra una pared al lado de la cual hay una caja de dimensiones $a \times b$ como en la figura 1.6. Formule el problema de encontrar la escalera mas corta posible.

9. Juan debe cortar n_i barras de longitud a_i, $i = 1, 2, \ldots, d$, de barras de una longitud fija L, $a_i \leq L$ para todo i. ¿Cuál es el numero minimo de barras que necesita? Encuentre una formulación precisa a este problema.

10. Un aeroplano esta volando a velocidad v con respecto al suelo en un campo de viento irrotacional acotado dado por $\nabla\varphi(x, y, z)$ y tal que $v > |\nabla\varphi|$. Comenzando y terminando en el mismo punto ¿Cuales son las distancias maximas y minimas que se pueden volar dado un intervalo de tiempo $[0, T]$? Plantee el problema asumiendo que v es constante, y la dirección de la velocidad esta a nuestra disposición. (Ayuda: Escriba una parametrización de la curva

$$\sigma(t) = (x(t), y(t), z(t))$$

¿Qué sabemos de x',y' y z' en terminos de v, la dirección de la velocidad y $\nabla\varphi$? Tenga presente que la longitud de esa curva esta dada por

$$\int_0^T |\sigma'|dt.)$$

11. Una cuerda esta colgada verticalmente en equilibrio de su extremo superior. (Figura 1.7). Esta se estira por acción de su propio peso y una masa constante W en su extremo inferior. El problema consiste en determinar la dristribución optima de la sección transversal a(x),$0 \leq x \leq L$, para minimizar la elongación total. La longitud sin estirar L, el volumen total, la densidad ρ y el modulo de Young E son constantes y conocidos.

    a) ¿Cuál es la restricción integral relacionada con el volumen total que la función $a(x)$ debe cumplir para ser admisible?

    b) Dado un $a(x)$, sea $y(x)$ la distancia medida desde el extremo superior a la cual se mueve la sección a una distancia x cuando la cuerda no ha sido estirada por el peso W. Asuma que se aplica la ley de Hooke. El ""strain $y'(x)$ en cada punto es directamente proporcional (con constante de proporcionalidad $\frac{1}{E}$) al "stress"en ese punto. Donde el "stress en x es la fuerza hacia abajo divida por el area transversal $a(x)$ ¿Escriba esta ley en forma de ecuación?

    c) ¿Comó se expresa el objetivo en terminos de y? ¿Hay alguna otra restricción que se deba imponer en y?

12. El problema del descenso mas lento a la luna puede ser formulado de la siguiente manera. Si $v(t)$ y $m(t)$ son la velocidad y la masa combinada de la nave espacial y combustible en el tiempo t. σ es la velocidad (constante) relativa de ejección del combustible, y g es la gravedad, entonces la ley de estado se escribe.

$$(m + dm)(v + dv) - dm(v + \sigma) - mv = mgdt$$

o de forma equivalente

$$\frac{dv}{dt} = g + \frac{\sigma}{m}\frac{dm}{dt}$$

Si la razon de la ejección por unidad de tiempo $-\frac{dm}{dt}$ puede ser controlada dentro de un intervalo $[0, \alpha]$, formule el problema de un aterrizaje suave en un tiempo minimo en terminos presisos.

13. Se requiere que un jet alcance un punto en el espacio en un minimo de tiempo desde el despegue. Asumiendo que la energia total (cinetica mas potencial mas (menos) combustible) es constante, el jet quema combustible a una taza constante maxima, y tiene velocidad cero al despegue, formule el problema de optimización correspondiente. (Ayuda: La ecuación de energia total nos lleva al postulado

$$v^2 + 2gy = at$$

donde $v = (x', y')$ es la velocidad, g es la aceleración de la gravedad y a ea la taza constante maxima a la cual se quema el combustible)

14. El relación con el problema de la construccón de un medio refractor optimo surge el siguiente problema:

$$\text{Maximizar} \qquad y(1)$$

sujeto a

$$y''(x) - F(x)y(x) = 0, \qquad y(0) = 1, \qquad y'(0) = 0$$
$$F \geq 0, \qquad \int_0^1 F(x)dx = M$$

Reformule este problema como un problema de control optimo con un funcinal integral como objetivo.

15. Un modelo agregado de crecimiento economico puede ser descrito mediante las siguientes ecuaciones

$$Y(t) = F(L(t), K(t)),$$
$$K''(t) + \mu K(t) = Y(t) - X(t), l \qquad \frac{L'(t)}{L(t)} = n$$

donde Y es la salida simple de la economia, unsando dos entradas, trabajo (L) y capital (K), X denota la cantidad de consumo, μ es la taza de depreciación, la variable t indica tiempo, y n es la taza constante a la que crece el trabajo. El objetivo de esta economia es maximizar la integral de bienestar

$$\int_0^\infty u(X(t)/L(t))e^{-\rho t}dt$$

donde ρ es el factor de descuento por tiempo. Trate de simplificar la formulación del problema tanto como sea posible.

16. En algunos casos los problemas de optimización pueden no adaptarse a alguna de las formas descritas en este capitulo, principalmente debido a que el criterio de optimización es mas desarrollado que aquellos vistos en este texto. Por ejemplo, una unidad de parada hidraulica (Figura 1.8), como aquellas utilizadas en la industria ferroviaria, crea una fuerza de amortiguamiento dada por:

$$F = c\frac{v^2}{a^2}, \qquad 0 \leq x \leq x_m$$

donde c es una constante, $v = v(x)$ es la velocidad del amortiguando, $a(x)$ es el area de un orificio que se le permite cambiar con el desplasamiento x, y x_m es el maximo desplazamiento permitido bajo las restricciones geometricas apropiadas. El diseño de tal unidad busca escoger $a(x)$ de tal manera que minimice la fuerza en un impacto dado de de una masa m a una valocidad v_0. Muestre que el optimo es obtenido cuando $a(x)^2$ varia linealmente con x (Ayuda: La formula del trabajo y la energia es

$$\frac{1}{2}mv^2 = \frac{1}{2}mv_0^2 - \int_0^x F(s)ds$$

¿Qué información nos daesta ecuación al final del impacto cuanto $v = 0$?

Capítulo 2

Programación Lineal

2.1. Introducción

La característica principal de un problema de programación lineal (PPL) es que todas las funciones involucradas, la función objetivo y aquellas expresando las restricciones deben ser lineales. La aprarición de una función no lineal, en la función objetivo o en las restricciones es suficiente para rechazar el problema como un PPL

Definición 2.1 (Forma general de un PPL) *Un PPL es un problema de optimización de la forma general*

$$Minimizar \qquad cx = \sum_i c_i x_i$$

suejto a:

$$\sum_i a_{ji} x_i \leq b_j, \qquad j = 1, \ldots, p,$$

$$\sum_i a_{ji} x_i \geq b_j, \qquad j = p+1, \ldots, q,$$

$$\sum_i a_{ji} x_i = b_j, \qquad j = q+1, \ldots, m,$$

donde c_i, b_j, a_{ji} son datos del problema. Dependiendo de los valores particuales de p y q podemos tener restricciones de desigualdad de un tipo o del otro asi como restricciones de igualdad.

Podemos enterder mas la estructura y características de un PPL estudiando un ejemplo simple.

Ejemplo 2.2 *Considere el PPL*

$$Maximizar \qquad x_1 - x_2$$

sujeto a:

$$x_1 + x_2 \leq 1, \qquad -x_1 + 2x_2 \leq 2,$$
$$x_1 \geq -1, \qquad -x_1 + 3x_2 \geq -3.$$

Es interesante notar que la forma del conjunto de puntos en el plano que satisface todos los requerimientos que imponen las restricciones: Cada desigualdad representa un "medio espacio" a un lado de la línea correspondiente a cambiar la desigualdad por una igualdad. Por lo tanto la intersección de los cuatro medio espacios sera la "región factible" para nuestro problema. Note que este conjunto tiene la forma de un polígono o un poliedro. Vease la figura 2.1.

Por otra parte,, el costo siendo lineal tiene curvas de nivel que son lineas rectas de ecuación $x_1 - x_2 = t$ donde t es una constante. A medida que t se mueve obtenemos lineas paralelas. La pregunta es que tan grande se puede volver t de tal modo que la línea de ecuación $x_1 - x_2 = t$ corta el polígono anterior en alguna parte. Gráficamente, no es difícil darse cuenta que el punto corresponde al vértice $(-1/2, 3/2)$, y el valor del máximo es 2.

Notese que independientemente de cual es el costo, mientras sea lineal, el valor óptimo siempre corresponderá a uno de los cuatro vertices del conjunto factible. Estos vertices juegan un papel crucial en el entendimiento de los PPL, como mas adelante veremos

Un PPL puede tomar varias formas. La forma inicial usualmente depende de la formulación particular del problema, o de la forma mas conveniente en la cual se pueden representar las restricciones. El hecho que todas las formulaciones corresponden al mismo problema de optimización nos permite fijar un formato de referencia, y referirnos luego a esta forma de cualquier problema particular para su análisis.

Definición 2.3 (Forma estándar de un PPL) *Un PPL en forma estándar es*

$$Minimizar \qquad cx \qquad bajo \qquad Ax = b, \qquad x \geq 0.$$

Por lo tanto los ingridientes de un PPL son:

1. una matriz de $m \times n$ con $n > m$ y generalmente n mucho mas grande que m;
2. un vector $b \in \mathbb{R}^m$
3. un vector $c \in \mathbb{R}^n$

Notese que cx es el producto interno de dos vectores c y x, mientras que Ax es el producto de la matriz A y el vector x. No se hara una distinción entre estas dos posiblidades ya que será claro en el contexto. Nuestro problema es por lo tanto encontrar el valor minimo que puede tomar el producto interno cx mientras x recorre todos los vectores factibles $x \in \mathbb{R}^n$ con componentes no negativas ($x \geq 0$) y que satisfagan la importante condición adicional $Ax = b$.

Tambien estamos interesados en el vector x (o en todos los vectores x) donde este valor mínimo se alcanza.

Hemos argumentado que cualquier PPL puede en principio ser transformado en la forma estándar. Es por lo tanto deseable que los lectores comprendan como se puede lograr esta transformación. Se procedera en tres pasos.

1. **Variables no restringidas en signo** Para variables no restringidas en signo usamos la descomposición de las variables en sus partes positivas y negativas de acuerdo a las indentidades:

$$x = x^+ - x^-, \qquad |x| = x^+ - x^-$$

donde

$$x^+ = \text{máx}\{0, x\} \geq 0, \qquad x^- = \text{máx}\{0, -x\} \geq 0$$

Lo que queremos decir con esta descomposición es que una variable x_i que no este restringida en signo puede ser escrita como la diferencia de dos nuevas variables que son no negativas.

$$x_i = x_i^{(1)} - x_i^{(2)}, \qquad x_i^{(1)}, x_i^{(2)} \geq 0.$$

2. **Transformando desigualdades en igualdades.** A menudo las restricciones son formuladas en forma de desigualdades. De hecho un PPL vendrá en la forma:

$$Minimizar \qquad cx \qquad bajo \qquad Ax \leq b, \qquad A'x' = b', \qquad x \geq 0.$$

Notese que usando la multiplicación por -1 podemos cambiar el sentido a una desigualdad. En esta situación el uso de "variables mudas" permite que pasemos de desigualdades a igualdades de la siguiente manera. Introduzca nuveas variables poniendo

$$y = b - Ax \geq 0.$$

Si ahora ponemos

$$X = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \widetilde{A} = \begin{pmatrix} A & \mathbf{1} \end{pmatrix}$$

Donde **1** Es la matríz identidad del tamaño apropiado, las restricciones de desigualdad ahora se escriben como

$$\widetilde{A}X = b$$

asi todas las restricciones estan ahora en la forma de igualdades, pero tenemos un numero mas grande de variables (uno mas por cada desigualdad)

3. **Transformando un max en un min.** Si el PPL nos pide un máximo en vez de un mínimo, debemos tener en cuenta que

$$\text{máx}(expresion) = -min(expresion)$$

o de manera mas explícita

$$max\{cx : Ax = b, x \geq 0\} = -min\{(-c)x : Ax = b, x \geq 0\}$$

Un ejemplo aclarara cualquier duda acerca de estas transformaciones.

Ejemplo 2.4 *Considere el PPL*

$$Maximizar \qquad 3x_1 - x_3$$

sujeto a

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_1 - x_2 - x_3 &\leq 1 \\ x_1 + x_3 &\geq 1 \\ x_1 \geq 0, \qquad x_2 &\geq 0 \end{aligned}$$

1. *Desde que hay variables que no estan restringidas en signo, debemos sutituir*

$$x_3 = y_1 - y_2 \qquad y_1 \geq 0, y_2 \geq 0$$

de manera que el probema cambie a:

$$Maximizar \qquad 3x_1 - y_1 + y_2$$

sujeto a

$$\begin{aligned} x_1 + x_2 + y_1 - y_2 &= 1 \\ x_1 - x_2 - y_1 + y_2 &\leq 1 \\ x_1 + y_1 - y_2 &\geq 1 \\ x_1 \geq 0, \qquad x_2 &\geq 0 \\ y_1 \geq 0, \qquad y_2 &\geq 0 \end{aligned}$$

2. *Usamos variables mudas de tal manera que las restricciones de desigualdades se transformen en igualdades: $z_1 \geq 0$ y $z_2 \geq 0$ se usan para transformar*

$$x_1 - x_2 - y_1 + y_2 \leq 1, \qquad x_1 + y_1 - y_2 \geq -1$$

respectivamente en

$$\begin{aligned} x_1 - x_2 - y_1 + y_2 + z_1 &= 1, \qquad z_1 = 0 \\ x_1 + y_1 - y_2 &\geq -1, \qquad z_2 = 0 \end{aligned}$$

El problema ahora tendrá la forma

$$Maximizar \qquad 3x_1 - y_1 + y_2$$

sujeto a

$$\begin{aligned} x_1 + x_2 + y_1 - y_2 &= 1 \\ x_1 - x_2 - y_1 + y_2 + z_1 &= 1 \\ x_1 + y_1 - y_2 - z_2 &= -1 \\ x_1 \geq 0, \qquad x_2 &\geq 0 \\ y_1 \geq 0, \qquad y_2 &\geq 0 \\ z_1 \geq 0, \qquad z_2 &\geq 0 \end{aligned}$$

3. *Finalmente, podemos cambiar facilmente el máximo a un mínimo*

$$Minimizar \qquad -3x_1 + y_1 - y_2$$

sujeto a

$$\begin{aligned} x_1 + x_2 + y_1 - y_2 &= 1 \\ x_1 - x_2 - y_1 + y_2 + z_1 &= 1 \\ x_1 + y_1 - y_2 - z_2 &= -1 \\ x_1 \geq 0, \qquad x_2 &\geq 0 \\ y_1 \geq 0, \qquad y_2 &\geq 0 \\ z_1 \geq 0, \qquad z_2 &\geq 0 \end{aligned}$$

teniendo en cuenta que una vez que se ha hallado este minimo, el correspondiente maximo sera el mismo pero con el signo cambiado.

Si unificamos la notación escribiendo

$$(X_1, X_2, X_3, X_4, X_5, X_6) = (x_1, x_2, y_1, y_2, z_1, z_2)$$

el problema obtendrá su forma estándar

$$Minimizar \qquad X_3 - X_4 - 3X_1$$

sujeto a

$$\begin{aligned} X_1 + X_2 + X_3 - X_4 &= 1 \\ X_1 - X_2 - X_3 + X_4 + X_5 &= 1 \\ X_1 + X_3 - X_4 - X_6 &= -1 \\ X &\geq 0 \end{aligned}$$

Una vez hemos resuelto este problema y tenemos una soloción optima X y el valor del mínimo m, la respuesta al PPL original se obtendra asi: El máximo es $-m$, y se optiene en el punto $(X_1, X_2, X_3 - X_4)$. O si lo prefieren, el valor del máximo sere el valor de la función costo original en la solución optima $(X_1, X_2, X_3 - X_4)$. Notese como las variables mudas no entran en la respuesta final, dado que son variables auxiliares.

Sobre la solución optima de un PPL, todas las situaciones pueden pasar:

1. El conjunto de vectores admisibles es el vacío.
2. No tiene solución, ya que el costo cx puede decrecer indefinidamente hacía $-\infty$ para vectores factibles x.
3. Puede admitir una sola solución unica, y esta es la situación mas deseable.
4. Puede tener muchas, de hecho infinitas soluciones optimas. De hecho es muy sencillo notar que si x_1 y x_2 son optimas, entonces cualquier combinacion convexa
$$tx_1 + (1-t)x_2, \qquad t \in [0,1]$$
es de nuevo una solución optima.

En la sección siguiente, veremos como resolver un PPL en su forma estándar con el método simplex. A pesar que los métodos de punto interior se estan volviendo más y más importantes en programación matematica, en ambas versiones lineal y no lineal, creemos que estos son el tema de un segundo curso programación matematica. El hecho es que el metodo simplex nos ayuda inmensamente a entender la estructura de la programación lineal asi como la dualidad.

2.2. El metodo simplex

Aquí miramos mas de cerca a un PPL en sufomr estándar, y describimos el metodo simplex, que es uno de los acercamientos mas exitosos para encontrar la solución optima de estos problemas. Concentremonos entonces en el problema de encontrar un vector x que resuelva

$$Minimizar \qquad cx$$

sujeto a

$$Ax = b, \qquad x \geq= 0$$

No hay restricción en asumir que el sistema $Ax = b$ tiene solución de lo contrario no habría vectores factibles. Ademas, si A no es una matriz de rango máximo podemos escojer una submatriz A' eliminando varias filas de A y los correspondientes componentes de b de tal manera que la nueva matriz A' tenga rango maximo. En este caso obtenemos un nuevo y equivalente PPL

$$Minimizar \qquad cx$$

sujeto a

$$A'x' = b' \qquad x \geq 0$$

donde b' es el subvector de b obtenido al eliminar las componentes correspondientes a las filas de A que hemos descartado anteriormente. Este nuevo PPL es equivalente al inicial en el sentido que ambos tienen las mismas soluciones

optimas, pero la matriz A' para el problema reducido es una matriz de rango máximo. Por lo tanto asumiremos sin perdida de generalidad que el rango de la matriz A de $m \times n$ es m (recuerde que $m \leq n$) y que el sistema lineal $Ax = b$ es solucionable.

Hay vectores factibles especiales que juegan un papel central en el metodo simplex. Estos son las soluciones del sistema lineal $Ax = b$ con componentes no negativas y al menos $n - m$ componentes nulas. De hecho todos estos puntos extremos o soluciones basicas como son tipicamente llamados, pueden, en principio, obtenerse resolviendo todos los sistemas cuadrados de tamaño $m \times m$ $Ax = b$ donde $n - m$ componentes de x son cero, y descartando aquellas soluciones con al menos una componente negativa. La estructura especial de un PPL nos permite concentrarnos en las soluciones basicas cuando estamos buscando soluciones óptimas.

Lema 2.5 *Si el PPL*

$$Minimizar \qquad cx$$

sujeto a

$$Ax = b \qquad x \geq 0$$

admite una solución óptima, entonces tambien hay una solución optima que es una solución basica.

Esto es bastante evidente si tenemos en cuenta que el conjunto factible el algun tipo de "poliedro" y por lo tanto los valores minimos o maximos de las funciones lineales deben tomarse en un vertice. Vease la figura 2.2 Y recuerdense los comentarios del ejemplo 2.2 Para la prueba de este lema, asumamos que x es una solución optima, con al menos $m + 1$ componentes estrictamente positivas, y sea d un vector no nulo en el kernel de A con la propiedad que $x_i = 0$ implica que $d_i = 0$. Si x tiene al menos $m + 1$ componentes estrictamente positivas, tal vector d siempre se puede encontrar (¿Por qué?).

Aseguramos que necesariamente $cd = 0$. Ya que de lo contrario si t es lo suficientemente pequeño de tal manera que el vector $x + td$ es factible (i.e., $x + td \geq 0$) y $tcd < 0$, entonces el costo del vector $x + td$ es estrictamente mas pequeño que el costo del vector x, que es imposible si x es optimo. Por lo tanto $cd = 0$, y los vectores $x + td$ son optimos mientras sean todos factibles. Todo lo que falta hacer es mover t de cero (en cualquier sentido positivo o negativo) hasta que alguna de las componentes de $x + td$ llegue a cero por primera vez. Para tal valor de t tenemos una solución optima con al menos un componente nulo mas que x. Este proceso puede ser repetido mientras el vector d no es el vector cero, i.e., hasta que la solución optima tiene al menos $n - m$ componentes nulas.

Como una consecuencia inmediata del lema 2.5, podemos encontrar una solución optima para un PPL mirando todas las soluciones del sistema $Ax = b$ con al menos $n - m$ ceros, descartando aquellas con componentes negativos, y

calculando el costo de las que quedan decidir el vector optimo. Este proceso nos llevara a una solución optima, pero el metodo simplex apunta a organizar estos calculos de tal manera que podamos llegar a una solución optima sin tener que pasar por un análisis exhaustivo de todos los puntos extremos. En algunos casos el metodo simplex pasa por todas las soluciones antes de encontrar una solución optima. Esta situación es sin embargo rara.

El metodo simplex (MS) comienza con un vector particular extremo x, el cual, despues de una permutación apropiada de los indices, puede ser escrito como:

$$x = \begin{pmatrix} x_B \\ 0 \end{pmatrix} \qquad x_B \in \mathbb{R}^m \qquad 0 \in \mathbb{R}^{n-m} \qquad x_B = 0$$

El paso básico iterativo consiste en poner uno de los componentes de x_B como cero (la variable saliente) y dejando que una componente nula de $0 \in \mathbb{R}^{n-m}$ (la variable entrante) se vuelva positiva. De esta manera nos hemos movido de un punto extremo a uno adyacente. La clave consiste en entender como hacer estas escogencias (las variables entrantes y salientes) de tal manera que reduzcamos el costo tanto como sea posible. Mas aun necesitamos un criterio para decidir cuando el costo no puede ser reducido y no se necesitan mas pasos iterativos. Mas precisamente.

Sea

$$x = \begin{pmatrix} x_B \\ 0 \end{pmatrix}, \qquad x_B \in \mathbb{R}^m \qquad 0 \in \mathbb{R}^{n-m} \qquad x_B \geq 0$$

un vector factible extremal?????. De la misma manera, la matriz A despues de la misma permutación de columnas puede ser descompuesta como

$$A = \begin{pmatrix} B & N \end{pmatrix}$$

La ecuación $Ax = b$ es equivalente a

$$\begin{pmatrix} B & N \end{pmatrix} \begin{pmatrix} x_B \\ 0 \end{pmatrix} = b, \qquad x_B = B^{-1}b.$$

El coste de este vector x es

$$cx = \begin{pmatrix} c_B & c_N \end{pmatrix} \begin{pmatrix} x_B \\ 0 \end{pmatrix} = c_B x_B = c_B B^{-1} b$$

El paso basico en el metodo simplex consiste en moverse a otro punto factible (adyacente) extremo de tal modo que el costo ha sido disminuido. El cambio desde $\begin{pmatrix} x_B & 0 \end{pmatrix}$ a $\begin{pmatrix} \bar{x_B} & x_N \end{pmatrix}$ donde x_N esta a nuestra disposición se hara si podemos asegurar tres requisitos:

1. $A\bar{x} = b$
2. $c\bar{x} < cx$

3. $\bar{x} \geq 0$

La primera condición nos fuerza a

$$\bar{x_B} = x_B - B^{-1}Nx_N$$

Lo cual es cierto, note que

$$\begin{pmatrix} B & N \end{pmatrix} \begin{pmatrix} \bar{x_B} \\ x_N \end{pmatrix} = b$$

implica

$$\bar{x_B} = B^{-1}(b - Nx_N) = x_B - B^{-1}Nx_N$$

En consecuencia el nuevo costo sera

$$\begin{pmatrix} c_B & c_N \end{pmatrix} \begin{pmatrix} x_B - \bar{B^{-1}}Nx_N \\ x_N \end{pmatrix} = c_B(x_B - B^{-1}Nx_N) + c_n x_N = (c_N - c_B B - 1N)x_N + c_B x_B$$

Notamos que el signo de

$$(c_N - c_B B^{-1}N)x_N$$

dira si hemos sido capaces de disminuir el costo moviendonos al nuevo vector

$$\begin{pmatrix} x_B - B^{-1}Nx_N & x_N \end{pmatrix}$$

El asi llamado vector de costos reducidos

$$r = c_N - c_B B^{-1}N$$

jugara in rol importante si podemos movernos a una nueva solución basica y reducir el costo. Dado que $x_N \geq 0$ (por el requerimiento 3) dos situaciones pueden ocurrir.

1. **Citerio de parada.** Si todas las componentes de r resultan no negativas, no hay manera de reducir el costo, y el punto extremal actual, es de hecho optimo. Hemos encontrado la solución a nuestro problema

2. **Paso iterativo** Si r tiene componentes no negativas, podemos en principio reducir el costo dejando que esas componentes de x_N se vuelvan positivas. Aunque, debemos realizar este cambio con precaución para asegurar que el vector

   $$x_B - B^{-1}Nx_N \tag{2.1}$$

   es factible. i.e este tiene solamente coordenadas no negativas. Si este no es el caso, aunque el costo sera menor en el vector

   $$\begin{pmatrix} x_B - B^{-1}Nx_N & x_N \end{pmatrix}$$

   este no sera factible y por lo tanto no admisible como una solución optima de el PPL. Debemos asegurar la no negatividad de todos los vectores extremos.

En vez de buscar soluciones mas generales de x_N el metodo simplex se enfoca en tomar $x_N = tv$, donde $t \geq 0$ es un vector de la base que tiene coordenadas nulas en todas partes excepto en una posición donde tiene un 1. Esto quiere decir que cambiaremos una componente a la vez. La componente escogida es precisamente la "variable entrante". ¿Cómo se selecciona esta variable? De acuerdo a nuestra discución previa estamos tratando de asegurar que el producto

$$rx_N = trv$$

sera tan negativo como se pueda. Dado que v es un vector de la base, rv es una componente de r, y por lo tanto v debe ser escogido como el vector de la base correspondiente al valor mas negativo de r. Una vez v ha sido seleccionado, tenemos que examinar

$$x_B - B^{-1}Nx_N = B^{-1}b - tB^{-1}Nv \tag{2.2}$$

para determinar la variable saliente. La idea es la siguiente. Cuando $t = 0$, estamos en nuestra solución basica x_B. ¿Qué puede pasar si t comienza a moverse desde cero a la parte positiva? En este punto pueden suceder tres situaciones que discutimos seguidamente.

1. **Solución no factible**. Lo pronto t se vuelve positiva el vector en (2.2) no es factible porque una de sus componentes es menor que cero. En este caso no podemos usar la variable escogida para reducir el costo, y debemos usar la siguiente variable negativa en r; o alternativamente podemos simplemente tomar esta variable como la variable saliente dado el hecho que el costo no disminuira. Usualmente se prefiere la segunda opción debido a su coherencia con el proceso.

2. **Variable Saliente.** Hay un valor limite t para el cual una de las coordenadas de (2.2) se vuelve cero por primera vez. Escogemos presisamente esta como la variable saliente saliente, y calculamos un nuevo punto extremo con un costo menor que el anterior.

3. **Ninguna solución.** Sin importar que tan grande se vuelve t, podemos seguir bajando el costo, y ninguna de las componentes de (2.2) llegara a cero. El problema no admite una solución optima porque podemos reducir el costo de manera indefinida.

El problema ahora consiste en como podemos decidir en cada situación particular en cual de los casos anteriores nos encontramos y proceder de forma correspondiente. Notese que cada expresión en (2.2) representa una linea recta en función de t. Las tres posibilidades estan dibujadas en la figura 2.3.

Asuma que hemos escogido la variable entrante indentificada con un vector de la base v. Procedemos de la siguiente manera.

1. Si una de las componentes nulas de $x_B = B^{-1}b$ corresponde a una componente positiva de $B^{-1}Nv$(diagrama 1 de la figura 2.3), entonces lo pronto t

se vuelve positivo esta coordenada sera menor que cero en (2.2), y el vector no sera factible. Podriamos recurrir a una variable entrante diferente (un vector v de la base distinto), que corresponde a aluna otra componente no negativa de r, si es posible. Si r no tiene mas componentes no negativas, ya tenemos la solución optima, y el metodo simplex se detiene. Alternativamente, y esta opción se prefiere tipicamente por coherencia, podemos considerar la razon de anulamiento como canditado para el proceso en 2.

2. Examine las razones de los vectores $B^{-1}b$ sobre $B^{-1}Nv$ componente por componente, y escoja como la variable saliente la correspondiente a la menor de estas razones entre las estrictamente positivas, incluyendo, como se menciono antes, las razones de anulamiento con denominador positivo. Estos podrian ser escogidos si estan presentes, dado que son mas pequeños que los estrictamente positivos. Ignore todos los cocientes sobre 0 incluyendo 0/0. Comience de nuevo el proceso con el nuevo vector extrema. Notese que estas razones corresponden a los valores de t cuando t intersecta el eje horizontal en el diagrama 2 de la figura 2.3.

3. Si no hay razones positivas, el PPL no admite una solución optima dado que el costo puede ser reducido infinitamente aumentando la variable entrante. Esto ocurre cuando todos los digramas son del tipo 3 en la figura 2.3.

Dado que el conjunto de vectores factibles es finito, despues de un numero finito de pasos, el metodo simplex nos lleva a una solución optima o a la conclusión que no hay un vector optimo. En algunas cirunstancias muy particulares, el metodo simplez puede entrar en un proceso cílico infinito. Esos casos son tan raros que no les prondremos atención. Un ejemplo simple es propuesto al final del capitulo.

En la practica los calculos pueden ser organizados de la siguiente manera algorítmica.

1. ***Inicialización.*** Encuentre una matriz cuadrada de tamaño $m \times m$ B de tal manera que la solución del sistema lineal $Bx_B = b$ es tal que $x_B \geq 0$.

2. ***Citerio de parada.*** Escriba

$$c = \begin{pmatrix} c_B & c_N \end{pmatrix}, \quad A = \begin{pmatrix} B & N \end{pmatrix}$$

Resuelva

$$z^T B = c_B$$

y mire al vector

$$r = c_N - z^T N$$

si $r \geq 0$ pare. Ya tenemos una solución optima. De lo contrario escoja la variable entrante como la componente mas negativa de r

3. ***Paso iterativo principal.*** Resuelva

$$Bw = y$$

donde y es la colomna entrante de N correspondiente a la variable entrante, y mire a las razones x_B/w componente a componente. Entre estas razones escoja aquellos con denominador positivo. Escoja como la variable saliente aquella correspondiente a la razon mas pequeña entre los seleccionades. Vaya al paso 2. Si no hay variable que escoger el problema no admite soluciones optimas.

Hemos tratado de reflejar el paso iterativo principal en la figura 2.4 !!!!!!!!!!!!Como se encuentra la matriz B para el siguiente paso??????

Para asegurar que nuestros lectores entiendan la estrategia en el metodo simplex, como se escogen las variables entrantes y salientes, y el criterio de parada, veremos varios ejemplos simples.

Ejemplo 2.6 *(Solución Unica)*

$$Minimizar \quad 3x_1 + x_2 + 9x_3 + x_4$$

sujeto a

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 4 \\ x_2 + x_3 - x_4 &= 2 \\ x_i &\geq 0 \end{aligned}$$

En este ejemplo particular,

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & -1 \end{pmatrix} \quad b = \begin{pmatrix} 4 \\ 2 \end{pmatrix} \quad c = \begin{pmatrix} 3 & 1 & 9 & 1 \end{pmatrix}$$

1. *Inicialización. Escoja*

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad N = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \quad c_B = \begin{pmatrix} 3 & 1 \end{pmatrix} \quad c_N = \begin{pmatrix} 9 & 1 \end{pmatrix}$$

2. *Revisar el criterio de parada. Es tribial encontrat*

$$x_B = b = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

de tal manera quie el vértice inicial es $\begin{pmatrix} 4 & 2 & 0 & 0 \end{pmatrix}$ *con costo 14. Por otra parte,*

$$z = c_B = \begin{pmatrix} 3 & 1 \end{pmatrix} \quad r = \begin{pmatrix} 9 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \end{pmatrix}$$

Dado que no todos los compontentes de r son no negativios, debemos pasar a travez del paso iterativo del metodo simplex.

3. *Paso iterativo. Escoja x_4 como la variable entrante, dado que esta es la asociada con la componente negativa de r. Mas aun,*

$$w = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \frac{x_B}{w}\{4, 2\}$$

de tal modo x_1 es la variable salient, siendo la que corresponde a la menor razon entre las que seleccionariamos (razones con denominadores positivos).

4. *Revisar el criterio de parada. Estos calculos nos llevan a la nueva escogencia*

$$B = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad c_B = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad c_N \begin{pmatrix} 3 & 9 \end{pmatrix}$$

Es facil encontrar

$$x_B = \begin{pmatrix} 6 \\ 4 \end{pmatrix}$$

y el nuevo vector extremo $\begin{pmatrix} 0 & 6 & 0 & 4 \end{pmatrix}$ con costo asociado 10. Los nuevos vectores z y r son

$$z = \begin{pmatrix} 2 & 1 \end{pmatrix}, \quad r = \begin{pmatrix} 3 & 9 \end{pmatrix} - \begin{pmatrix} 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \end{pmatrix}$$

Dado que todas las comonentes de r son no negativas hemos acabado nuestra busqueda. El costo minimo es 10, y se toma en el vector $\begin{pmatrix} 0 & 6 & 0 & 4 \end{pmatrix}$.

Ejemplo 2.7 *(Ejemplo Degenerado)*

$$\textit{Mínimizar} \quad 3x_1 + 2x_3 + 9x_3 + x_4$$

sujeto a

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 0 \\ x_2 + x_3 - x_4 &= 2 \\ x_i &\geq 0 \end{aligned}$$

En este caso particular

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad c = \begin{pmatrix} 3 & 1 & 9 & 1 \end{pmatrix}$$

1. *Inicialización. Escoja*

$$B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad n = \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix}, c_B = \begin{pmatrix} 3 & 1 \end{pmatrix}, \quad c_N = \begin{pmatrix} 9 & 1 \end{pmatrix}$$

2. *Revisar el criterio de parada. Es trivial encontrar*

$$x_B = b = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

por lo que el vector inicial es $\begin{pmatrix} 0 & 2 & 0 & 0 \end{pmatrix}$ *con costo 2. Por otra parte,*

$$z = c_B = \begin{pmatrix} 3 & 1 \end{pmatrix}, \quad r = \begin{pmatrix} 9 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \end{pmatrix}$$

Dado que no todos los componentes de r son no negativos, debemos parar el paso iterativo del metodo simplex.

3. *Paso iterativo. Escoja a* x_4 *como la variable entrante, dado que esta asociada con la componente negativa de r. Ademas*

$$w = \begin{pmatrix} 1 & -1 \end{pmatrix}, \quad \frac{x_B}{w} = 0, -2$$

entonces x_1 *es la variable saliente, siendo aquella la correspondiente a la minima razon entre las que seleccionariamos (razones con denominador positivo). Podemos predecir sin embargo, que porque nuestra unica opcion es una razon nula no seremos capaz de disminuir el costo a pesar de atravezar el paso iterativo del metodo simplex. En otras palabras, el vector* $\begin{pmatrix} 0 & 2 & 0 & 0 \end{pmatrix}$ *ya es una solución optima. Dado que para esta solución no se cumple el criterio de parada para nuestra escogencia original de B (r tiene coordenadas negativas), debemos, por coherencia, pasar por el paso iterativo del metodo simplex.*

4. *Revisar el criterio de parada. La nueva escogencia*

$$B = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad c_B = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad c_N = \begin{pmatrix} 3 & 9 \end{pmatrix}$$

nos lleva a

$$x_B = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

y el nuevo vector externo es de nuevo $(0, 2, 0, 0)$ *con costo asociado 2. Los vectores z y r son*

$$z = \begin{pmatrix} 2 & 1 \end{pmatrix}, \quad r = \begin{pmatrix} 3 & 9 \end{pmatrix} - \begin{pmatrix} 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 4 \end{pmatrix}$$

Dado que todas las componentes de r son no negativas, hemos terminado nuestra busqueda como lo habiamos anticipado: El costo mínimo es 2 y se toma en el vector $(0, 2, 0, 0)$.

Ejemplo 2.8 *(Sin solución)*

$$\textit{Mínimizar} \quad -3x_1 + x_2 + 9x_3 + x_4$$

sujeto a

$$\begin{aligned} x_1 - 2x_3 - x_4 &= -2 \\ x_2 + x_3 - x_4 &= 2 \\ x_i &\geq 0. \end{aligned}$$

En este caso,

$$A = \begin{pmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & 1 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \quad c = \begin{pmatrix} -3 & 1 & 9 & 1 \end{pmatrix}.$$

1. *Inicialización. Si escogiéramos*

$$B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad N = \begin{pmatrix} -2 & -1 \\ 1 & -1 \end{pmatrix}, \quad c_b = \begin{pmatrix} -3 & 1 \end{pmatrix}, \quad c_n = \begin{pmatrix} 9 & 1 \end{pmatrix},$$

entonces obtendriamos

$$x_B = b = \begin{pmatrix} -2 \\ 2 \end{pmatrix}$$

que no es un vector factible dado que tiene una coordenada negativa. Tomemos en cambio la segunda y la cuarta columna de A.

$$B = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}, \quad c_B = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad c_N = \begin{pmatrix} -3 & 9 \end{pmatrix}$$

2. *Revisando el criterio de parada. Es facil encontrar*

$$x_B = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

de tal modo qye el vector inicial es $(0, 4, 0, 2)$ *con costo 6. Por otro lado,*

$$z = \begin{pmatrix} -2 & 1 \end{pmatrix}, \quad r = \begin{pmatrix} -3 & 9 \end{pmatrix} - \begin{pmatrix} -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 4 \end{pmatrix}$$

Dado que no todos los componentes de r son no negativo, debemos pasar a traves del paso iterativo del metodo simplex.

3. *Paso iterativo. Escoja* x_1 *como la variable entrante, dado que esta es la que esta asociada con la componente negativa de r. Por consiguiente.*

$$w = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \frac{x_B}{w} = \{-4, -2\}$$

En esta situación no tenemos opción para la variable saliente ya que no hay denominadaor positivo. Esto quiere decir que el PPL no admite una

solucíon óptima, i.e El costo puede ser reducido indefinidamente. Esto puede ser facilmente comprobado considerando los vectores factibles

$$\begin{pmatrix} t-2 \\ t+2 \\ 0 \\ t \end{pmatrix}, \qquad t \geq 0.$$

El costo asociado con estos puntos es $8-t$, que puede ser claramente enviado a $-\infty$ haciendo t lo suficientemente grande.

Ejemplo 2.9 *(Soluciónes multiples)*

$$\textit{Mínimizar} \quad 3x_1 + 2x_2 + 8x_3 + x_4$$

sujeto a

$$\begin{aligned} x_1 - 2x_3 - x_4 &= -2 \\ x_2 + x_3 - x_4 &= 2 \\ x_i &\geq 0 \end{aligned}$$

Para argumentar que hay infinitas soluciones optimas para este PPL, usaremos las restricciones de igualdad para "despejar" x_1 y x_2 y devolver estas expresiones a la función objetivo. Especificamente.

$$x_1 = 2x_3 + x_4 - 2 \geq 0, \quad x_2 = -x_3 + x_4 + 2 \geq 0.$$

Y la función de costo se converte en

$$6(2x_3 + x_4) - 2$$

Dado que la primera restricción dice

$$2x_3 + x_4 \geq 2,$$

Es claro que el, costo mínimo se alcanzara cuando

$$2x_3 + x_4 = 2,$$

Tenemos dos soluciónes basicas $(0,1,1,0)$ y $(0,4,0,2)$. Cualquier combinación convexa tambien sera una solución optima

$$t\begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix} + (1-t)\begin{pmatrix} 0 & 4 & 0 & 2 \end{pmatrix}, \quad t \in [0,1]$$

Creemos que es elemental entender la forma en que funciona el metodo simplex despues de varios ejemplos. Sin embargo los calculos pueden ser organizados en tablas para facilitar todo el proceso sin tener que explicitamente escribir los diferentes pasos como hemos hecho hasta ahora. Trataremos esos detalles practicos en una sección posterior.

2.3. Dualidad

La dualidad es un concepto que vincula los siguientes dos PPL

$$\begin{array}{lllll} \text{Mínimizar} & cx & \text{sujeto a} & Ax \geq b, & x \geq 0; \\ \text{Máximizar} & yb & \text{sujeto a} & yA \leq c, & y \geq 0; \end{array}$$

Iderntificaremos el primer problema como el primario, y el segundo como su dual asociado. Notese como los mismos elementos, la matriz A y los vectores b y c, determinan ambos problemas.

Definición 2.10 *(Problema dual) El problema dual de el PPL*

$$\textit{Mínimizar} \quad cx \quad \textit{sujeto a} \quad Ax \geq b, \quad x \geq 0$$

es el PPL

$$\textit{Máximizar} \quad yb \quad \textit{sujeto a} \quad yA \leq c, \quad y \geq 0$$

Aunque este formato no es el que hemos utilizado en nuestra discusión del metodo simplex, nos permite ver de forma mas transparente que el dual del dual es es el primario. Esto es de hecho bastante sencillo de verificar transformando minimos en maximos y cambiando el sentido de las desigualdades usando apropiadamente signos menos (se deja esto al lector).

Si el problema primario es formulado en la forma estandar ¿Cúal es su versión dual? Contestar esta pregunta es un ejercicio elemental que consiste en escribir un PPL en su forma estandar en el formato anterior, aplicar dualidad, y entonces tratar de simplificar la forma final del dual. De hecho, lo unico que tenemos que hacer es poner

$$Ax = b \quad \text{es equivalente a} \quad Ax \geq b, \quad -Ax \geq -b,$$

asi si escribimos

$$\bar{A} = \begin{pmatrix} A & -A \end{pmatrix}, \quad \bar{b} = \begin{pmatrix} b & -b \end{pmatrix},$$

nuestro PPL inicial es

$$\text{Minimizar} \quad cx \quad \text{bajo} \quad \bar{A}x \geq \bar{b}, \quad x \geq 0.$$

Por lo tanto su dual tendra la forma

$$\text{Máximizar} \quad \bar{y}\bar{b} \quad \text{sujeto a} \quad \bar{y}\bar{A} \leq c, \quad \bar{y} \geq 0.$$

Si ahora tratamos de simplificar la formulación del problema poniendo

$$\bar{y} = \begin{pmatrix} y^{(1)} & y^{(2)} \end{pmatrix}$$

llegamos a

$$\text{Máximizar} \quad \begin{pmatrix} y^{(1)} & y^{(2)} \end{pmatrix} b \quad \text{sujeto a} \quad \begin{pmatrix} y^{(1)} & y^{(2)} \end{pmatrix} A \leq c, \quad \bar{y} \geq 0.$$

y poniendo $y = y^{(1)} - y^{(2)}$, no hay restricción en el signo de y, y tenemos

$$\text{Máximizar} \quad yb \quad \text{sujeto a} \quad yA \leq c.$$

que es la forma del dual cuando el primario esta dado en la forma estandar.

Lema 2.11 *(Problema dual en forma estandar) Si el problema primario es*

$$\textit{Mínimizar} \quad cx \quad \textit{bajo} \quad Ax = b, \quad x \geq 0, \tag{2.3}$$

su dual es

$$\textit{Máximizar} \quad yb \quad \textit{sujeto a} \quad yA \leq c.$$

Antes de proseguir en analizar la dualidad de forma mas formal, vale la pena motivar su análisis dando una interpretación del significado de la relación del primario y su dual. De hecho es interesante notar que son formas diferentes pero equivalentes de mirar el mismo problema de fondo. Vamos a enfatizar este punto describiendo un típico PPL relacionado con redes. Las implicaciones practicas de la dualidad estan frecuentemente atadas al problema de fondo detras de un PPL formal. Aquí restringimos nuestra atención a revisar formalmente la equivalencia entre del primario y su dual sin poner mucha atención a otras implicaciones. Un análisis y entendimiento completo de estos seria requerido en sitaciones realistas.

Ejemplo 2.12 *Deseamos enviar cierto producto desde el nodo A al nodo D en la red simplificada de la figura 2.5. Como se puede ver, tenemos cinco posibles canales con costos asociados dados en la misma figura. Su usamos las variables x_{PQ} para denotar la fracción del producto transferido a traves del canal PQ, debemos mínimizar el costo total*

$$2x_{AB} + 3x_{AC} + x_{BC} + 4x_{BD} + 2x_{CD}$$

sujeto a las restricciones

$$x_{AB} = x_{BC} + x_{BD}$$
(parte del producto no se pierde en el nodo B),

$$x_{AC} + x_{BX} = x_{CD}$$
(parte del producto no se pierde en el nodo C),

$$x_{BD} + x_{CD} = 1$$
(todo el producto llega al nodo D)
$$x_{AB}, x_{AC}, x_{BC}, x_{BD}, x_{CD} \geq 0.$$

Notese que como consequencia de estas restricciones es facil obtener

$$x_{AB} + x_{AC} = 1$$

asi que la totalidad del producto sale del nodo A. Esta es la formulación primaria del problema.

Tambien podemos pensar en terminos de precios por unidad de producto en los diferentes nodos de la red y_A, y_B, y_C, y_D y considerar las diferencias entre

estos precios como la ganancia cuando se usa algun canal particualar. En esta situación estamos buscando el máximo para $y_D - y_A$, la ganancia en transferir el producto de A a D. Las ganancias por los cinco canales serán

$$y_B - y_A, \quad y_C - y_A, \quad y_C - y_B, \quad y_D - y_B, \quad y_D - y_C.$$

Si tomamos como regla de normalización $y_A = 0$, entonces debemos pedir que estas ganancias no exedan los precios por el uso de cada canal:

$$y_B - y_A = y_B \leq 2, \quad y_C - y_A = y_C \leq 3, \quad y_C - y_B \leq 1,$$
$$y_D - y_B \leq 4, \quad y_D - y_C \leq 2.$$

Esta seria la formulación dual del problema.

De alguna forma sospechamos que estos dos problemas deben ser equivalentes y que sus soluciones optimas deben estar relacionadas la una con la otra. De hecho ese es el caso. La conección es precisamente la dualidad. Con los elementos

$$c = \begin{pmatrix} 2 & 3 & 1 & 4 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$
$$A = \begin{pmatrix} 1 & 0 & -1 & -1 & 0 \\ 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

estos dos problemas se pueden formular de manera compacta como

$$\textit{Mínimizar} \quad cx \quad \textit{bajo} \quad Ax = b, \quad x \geq 0,$$
$$\textit{Máximizar} \quad yb \quad \textit{bajo} \quad yA \leq c.$$

donde

$$x = \begin{pmatrix} x_{AB} & x_{AC} & x_{BC} & x_{BD} & x_{CD} \end{pmatrix}, \quad y = \begin{pmatrix} y_B & y_C & y_D \end{pmatrix}$$

Estqa es precisamente la forma de un primario y su dual.

A continuación describimos brevemente en dos pasos la relación entre las soluciones del problema primario (P) y su dual (D), donde asumimos que (P) esta dado en forma estándar.

Lema 2.13 *(Dualidad debil) Si x y y son factibles para (P) y (D) respectivamente, entonces*

$$yb \leq cx$$

Mas aun si hay igualdad

$$yb = cx$$

entonces x es una solución optima para (P) y y para (D).

La prueba es bastante sencilla, de

$$Ax = b, \quad x \geq 0, \quad yA \leq c,$$

tenemos

$$yb = yAx \leq cx.$$

En particular tenemos

$$\text{máx}\{yb : yA \leq c\} \leq \text{mín}\{cx : Ax = b, x \geq 0\}.$$

Si $yb = cx$ este numero debe ser al mismo tiempo el máximo y el mínimo anterior, y esto implica que y es optimo para (D) y x para (P).

Este resultado tambien nos informa que en los casos degenerados, en los que el máximo para el dual es $+\infty$ o el minimo del primario es $-\infty$, el otro problema no tiene vectores factibles.

De esta observación se sigue el teorema de dualidad completo.

Teorema 2.14 *(Teorema de dualidad) O ambos problemas (P) y (D) son solubles simultaneamente, o uno de los dos es degenerado en el sentido que no admite vectores factibles.*

Lo que dice esta afirmación es que si x es óptimo para (P), entonces existe un vector óptimo y para (D), y el valor comun $yb = cx$ es al mismo tiempo el mínimo para (P) y el máximo para (D). Conversamente, si y es optimo para (D), existe un vector optimo x para (P) con el valor comun $yb = cx$ siendo al mismo tiempo el mínimo y el máximo.

La prueba se basa en nuestra discución previa del metodo simplex. Si x es optimo para (P), entonces

$$x = \begin{pmatrix} x_B & 0 \end{pmatrix}, \quad x_B = B^{-1}b, \quad cx = c_B B^{-1} b,$$
$$r = c_N - c_B B^{-1} N \geq 0 \text{ (Criterio de parada)},$$

donde $c = \begin{pmatrix} c_B & c_N \end{pmatrix}$. Si examinamos $y = c_B B^{-1}$, pasa que

$$yA = c_B B^{-1} \begin{pmatrix} B & N \end{pmatrix} = \begin{pmatrix} c_B & c_B B^{-1} N \end{pmatrix} \leq \begin{pmatrix} c_B & c_N \end{pmatrix} = c,$$

asi que y es admisible para (D). Por otro lado, y es tal que

$$cx = c_B B^{-1} b = yb$$

Por el principio de dualidad debil, x y y deben ser óptimos.

El hecho que el dual del dual es el primario nos permite pasar de una solución optima para (D) a una solución optima para (P).

Note como se ha obtenido una solución para el dual de la solución optima del primario: Si $x = \begin{pmatrix} x_B & 0 \end{pmatrix}$ es óptimo para P con

$$x_B = B^{-1}b, \quad c = \begin{pmatrix} c_B & c_N \end{pmatrix}$$

entonces $y = c_B B^{-1}$ es óptimo para (D). Regresaremos a esta observación despues.

Es importante enfatizar la información en el primario dada por la solución optima del dual. Una de estas interpretaciones viene directamente de el teorema de dualidad, y refiere como cambios al vector b afectan el valor optimo del primario. Este es un tema de gran importancia practica, dado que tambien estamos interesados en conocer que tan buenos son los cambios en el vector b. Las restricciones que impone b estan tipicamente relacionadas con restricciones como capacidades de producción e inversiones totales, y nos gustaria conocer si hacer esos cambios seria rentable. Si $M(b)$ es la dependencia del valor mínimo de PPL de b, estamos buscando las derivadas parciales ∇M. Estos sol llamados parametros de sensibilidad, precios sombra asi como precios duales. Por el teorema de dualidad tenemos,

$$M(b) = y(b)$$

Donde y es la solución optima del dual. Intuitivamente,

$$\nabla M(b) = y$$

y esto es usualmente interpretado diciendo que el dual da la taza de cambio de la solución optima del primario cuando el vector b de restricciones cambia. Es por lo tanto importante saber la solución optima de un PPL asi como la solución optima del dual. Para ser rigurosos, el cálculo anterior de ∇M no ha sido justificado, dado que la solución optima del dual depende de b. Pero dado que el resultado es esencialmente correcto y plausible, no insistiremos en este punto.

2.4. Algunos asuntos practicos

En las secciones anteriores, hemos estado interesandos en la comprensión de la estructura de un PPL y del mecanismo estándar para resolverlo con el metodo simplex. Hay sin embargo un numero de asuntos de importancia practica. Trataremos en esta sección tres de estos temas.

1. Como inicializar el metodo simplex desde un punto de vista practico.
2. Como organizar los calculos en una forma eficiente a travez de tablas.
3. Como la solución optima del dual se puede encontrar rapidamente de la solución del primario.

La importancia de de dichos asuntos es de valor relativo, dado que tan pronto como el numero de variables involucradas en un PPL pasa de unas cuantas, paquetes computacionales deben ser usados para encontrar soluciones optimas in periodos razonables de tiempo.

En nuestra discusión del metodo simplex, no hemos indicado como encontrar una primera escogencia factible para la matriz B. Esto es seleccionar m columnas

entre las n columnas de A de tal manera que la solución del sistema lineal $Bx = b$ es tal que $x \geq 0$. En algunos casos, hacer esto directamente puede ser una tarea bastante tediosa. Lo que nos gustaria hacer es describir un mecanismo mas o menos eficiente que llevara a una submatriz factible B sin pasar por una enumeración exhaustiva de todas las posibilidades, lo que seria despues de todo resolver el PPL a fuerza bruta. Describiremos dos formas diferentes de encontrar tal inicialización.

El primero depende de un PPL auxiliar con inicialización trivial, cuya solución optima nos dira como seleccionar la matriz inicial factible B. El problema auxiliar es

$$\text{Mínimizar} \quad \sum_i \bar{x}_i$$

sujeto a

$$\begin{pmatrix} \bar{A} & \mathbf{1} \end{pmatrix} X = b, \quad X \geq 0.$$

donde $X = \begin{pmatrix} x & \bar{x} \end{pmatrix}$ y $\bar{A}$ y $\bar{b}$ son tales que el sistema $\bar{A}x = \bar{b}$ es equivalente al sistema $Ax = b$ pero $\bar{b} \geq 0$. Esto se puede hacer simplemente multiplicando -1 aquellas restricciones asociadas a componentes negativas de b. Note que un vector extremo factible valido para este problema es $\begin{pmatrix} 0 & \bar{b} \end{pmatrix}$.

Afirmamos que si el PPL inicial admite una solución óptima x, entonces el mínimo para el problema auxiliar es 0, y se logra en $X = \begin{pmatrix} x & 0 \end{pmatrix}$. Esto es muy facil de verificar y se le deja al lector como ejercicio. En cosecuencia si resolvemos este problema auxiliar con el vector inicial $\begin{pmatrix} 0 & \bar{b} \end{pmatrix}$ por el metodo simplex, la solución optima encontrada sera de la forma $X = \begin{pmatrix} x & 0 \end{pmatrix}$, donde x tiene como máximo m componentes no nulas. Observe que el problema auxiliar tiene el mismo valor de m. Las componentes positivas de este vector x indicaran cuales columnas deben ser escogidas para un punto inicual factible para nuestro PPL. Si el numero de tales componentes positivas es estrictamente menor que el numero de columnas que seleccionariamos, las columnas que quedan se pueden seleccionar de forma arbitraria, mientras se mantengan como un conjunto linealmente independiente de vectores. Para clarificar el mecanismo que lleva a una inicialización factible de cualquier PPL, consideremos el siguiente ejemplo.

Ejemplo 2.15 *Estamos interesados en encontrtar una inicialización valida para el PPL*

$$\textit{Minimizar} \quad 2x_1 + 3x_2 + x_3$$

sujeto a

$$\begin{aligned} x_1 + x_2 + 2x_3 + x_4 &= 500 \\ x_1 + x_2 + x_3 - x_4 &= 500 \\ x_1 + 2x_2 + 2x_3 &= 600 \\ x_i &\geq 0 \end{aligned}$$

En este ejemplo particular,

$$A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 2 & 1 & -1 \\ 1 & 2 & 2 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 500 \\ 500 \\ 600 \end{pmatrix}, \quad c = \begin{pmatrix} 2 & 3 & 1 & 0 \end{pmatrix}.$$

El asunto aquí es como escoger la matriz inicial B para inicializar el metodo simplex. En este caso simple tenemos cuatro posiblidades que corresponde a escoger tres diferentes columnos de un conjunto de cuatro. Ciertamente podriamos pasar por todas esas posibilidades y escoger la primera solución basica con coordenadas no negativas. Como hemos mencionado antes, esto es equivalente a resolver el PPL a traves de un análisis exahustivo de todas las soluciones basicas. Cuando la dimensión del problema es grande, este metodo enumerativo no es admisible, y el mecanismo descrito para inicializar el metodo simplex se vuelve interesante. En nuestro ejemplo este enfoque equivaldria a considerar el siguiente PPL asociado con los datos.

$$\bar{A} = \begin{pmatrix} 1 & 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & -1 & 0 & 1 & 0 \\ 1 & 2 & 2 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \bar{b} = \begin{pmatrix} 500 \\ 500 \\ 600 \end{pmatrix}$$
$$c = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}$$

La inicialización para el metodo simplex para resolver para resolver este problema es escoger la matriz B como la matriz identidad par las ultimas tres columnas de A. Las componentes no nulas de la solución óptima para este problema encontrada usando el metodo simplex indicaran una inicialización para nuestro problema original. En este caso, aplicando el metodo simplex obtenemos la solución óptima

$$\begin{pmatrix} 200 & 100 & 100 & 0 & 0 & 0 & 0 \end{pmatrix}$$

y esto nos indica que la matriz inicial B, hecha de las tres primeras columnas de A es una inicialización valida para el metodo simples para nuestro PPL original.

El segundo enfoque al problema de la inicialización no requiere considerar un PPL auxiliar. En cambio se basa en transformar el PPL en un PPL equicalente para el cual la inicialización es trivial. En vez de mencionar un resultado formal, discutiremos esta transformación intuitivamente. Considere un tipico PPL,

$$\text{Mínimizar} \quad cx \quad \text{bajo} \quad Ax = b, \quad x \geq 0,$$

donde $b \geq 0$ (multiplicando por -1 aquellas ecuaciones que corresponden a componentes negativas de b). Introduscamos nuevas variables $y \in \mathbb{R}^m$ y estudiemos el nuevo PPL

$$\text{Mínimizar} \quad cx + dy \quad \text{bajo} \quad Ax + y = b, \quad x \geq 0, \quad y \geq 0,$$

donde el vector $d \in \mathbb{R}^m$ se asume que tiene componentes no espacificadas muy grandes. El punto es que $\begin{pmatrix} 0 & b \end{pmatrix}$ es una inicialización valida para el segundo PPL, y dado que las componentes de d son lo suficientemente grandes, la solución óptima será de la forma $\begin{pmatrix} x & 0 \end{pmatrix}$, donde x es una solución óptima para nuestro PPL inicial. La primera afirmación es trivial. La segunda es plausible, dado que los componentes de d son muy grandes y estamos buscando mínimizar la suma $cx + dy$ con $y \geq 0$, vemos que las soluciones óptimas requieren que $y = 0$, y por lo tanto regresamos a nuestro PPL original. La desventaja de este procedimiento

es que para resolver este PPL transformado debemos trabajar simbolicamente con el vector d, o de lo contrario debemos asignar a d valores muy altos. Veremos un ejemplo despues.

Los calculos involucrados en el metodo simplex estan normalmente organizados el forma de tablas que reflejan los diferentes pasos que hemos descrito en la sección 2.2. Dado que el metodo simplex procede cambiando submatrices B de la matriz original A, y renombrando columnas de tal manera que las primeras m columnas corresponden a la matriz B, es muy importante no perderse con tal reorganización y mantener un registro de la enumeración inicial de las columnas independientemente de la posición que ocupan en los pasos posteriores. Cada una de estas tablas tiene la estuctura

$$\begin{array}{cc} A & b \\ \\ c & d \end{array}$$

donde d es el valor indicando el costo, cambiado en signo, de los diferentes soluciónes basicas factibles por las cuales pasa el metodo simplex. En cada una de estas tablas se deben hacer los siguientes calculos sucesivamente.

1. Escoger las columnas correspondientes a la siguiente submatriz B y ponerlas en las primeras m columnas de la tabla,y usando transformaciones basicas del algebra lineal transformar la matriz B en la matriz identidad (una matriz triangular superior o inferior no es suficiente). No olvide mantener un registro de las columnas de B.

2. Transforme de la misma manera las componentes de c ultima columna de tal manera que aquellos que correspondan a las columnas de B sean nulos.

3. Si los componentes que quedan de c son no negativos (criterio de parada), una (la) solución óptima se encuentra resolviendo el sistema lineal $Bx = b$ con la submatriz actual B y el termino independiente b, y poniendo cero en aquellas componentes que no esten en B (es por esto que es importante mantener un registro de cuales columnas son parte de las submatrices B). Si hay algunas componentes negativas, seleccionamos la columna entrante como aquella con la minima componente en c.

4. Examine las razones de b sobre la columna entrante componente por componente. Si no hay denominador positivo el problema no tiene solución óptima, de lo contrario escoja como la columna saliente aquella asociada con la mínima razon no negativa entre las seleccionadas. Regrese al paso 1 hasta que se satisfaga el criterio de parada, o lleguemos a la conclusión que no existe una solución óptima.

En vez de insistir en clarificar estos puntos que reflejan fielmente aquellos descritos en la sección 2.2, proponemos examinar un ejemplo concreto.

Ejemplo 2.16 *Queremos mínimizar* $-3x_1 - 5x_2$ *bajo las restricciones* $x \geq 0$ *y*

$$3x_1 + 2x_2 + x_3 = 18, \quad x_1 + x_4 = 4 \quad x_2 + x_5 = 0$$

La tabla inicial para este problema es

$$\begin{array}{ccccccc} x_1 & x_2 & x_3 & x_4 & x_5 & \\ 3 & 2 & 1 & 0 & 0 & 18 \\ 1 & 0 & 0 & 1 & 0 & 4 \\ 0 & 1 & 0 & 0 & 1 & 6 \\ & & & & & \\ -3 & -5 & 0 & 0 & 0 & 0 \end{array}$$

Si escogemos las columnas o variables 3, 4 *y* 5 *para hacer la matriz* B*, encontramos que el vertice* $\begin{pmatrix} 0 & 0 & 18 & 4 & 6 \end{pmatrix}$ *es factible. Si reorganizamos las tres columnas seleccionadas como las tres primeras columnas obtenemos la tabla*

$$\begin{array}{ccccccc} x_3 & x_4 & x_5 & x_1 & x_2 & \\ 1 & 0 & 0 & 3 & 2 & 18 \\ 0 & 1 & 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 0 & 1 & 6 \\ & & & & & \\ 0 & 0 & 0 & -3 & -5 & 0 \end{array}$$

Dado que en este caso particular la matriz B *ya es la identidad, no se requieren mas calculos en la tabla para estre proposito. Por otro lado, las componentes de* c *que no corresponden a las columnas de* B *son ambas negativas(*-3 *y* -5*), dado que el criterio de parada no se cumple, debemos transformar la tabla de acuerdo al paso principlal del metodo simplex. La variable entrante será* x_2 *(asociada con* -5 *en* c*). Para determinar la variable saliente, debemos examinar las razones* $18/2$ *y* $6/1$ *(*$4/0$ *se descarta porque tiene denominador nulo) y seleccionamos a la tercera razon* 6 *como la mas pequea. Segun esto la tercera columna de* B *(correspondiente a* x_5 *es la columna saliente. Despues que estas dos variables son intercambiadas la tabla se ve así*

$$\begin{array}{ccccccc} x_3 & x_4 & x_2 & x_1 & x_5 & \\ 1 & 0 & 2 & 3 & 0 & 18 \\ 0 & 1 & 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 0 & 1 & 6 \\ & & & & & \\ 0 & 0 & -5 & -3 & 0 & 0 \end{array}$$

Con esta tabla debemos obtener por medio de transformaciones elementales la matriz identidad en las primeras tres columnas, y el vector nulo en las componentes de c*. En este caso particular, estos dos objetivos se logran cambiando la primera fila por ella misma menos dos veces la tercera , y reemplazando la cuarta por si misma mas cinco veces la tercera. Despues de estos cambios llegamos a*

$$\begin{array}{ccccccc} x_3 & x_4 & x_2 & x_1 & x_5 & \\ 1 & 0 & 0 & 3 & -2 & 6 \\ 0 & 1 & 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 0 & 1 & 6 \\ & & & & & \\ 0 & 0 & 0 & -3 & 5 & 30 \end{array}$$

De nuevo miramos a las componentes no nulas de c y seleccionamos la mínima (negativa) como la variable entrante (x_1). Para escoger la variable saliente, examinamos las razones $6/3$ *y* $4/1$*, y escogemos la mas pequea entre las positivas, asoiada en este caso con* x_3*. Estos cambios llevan a*

$$\begin{array}{ccccc|c} x_1 & x_4 & x_2 & x_3 & x_5 & \\ 3 & 0 & 0 & 1 & -2 & 6 \\ 1 & 1 & 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 1 & 6 \\ \hline -3 & 0 & 0 & 0 & 5 & 30 \end{array}$$

Como antes buscamos la matriz identidad en las primeras tres columnas, y el vector nulo en las tres componentes de c. La nueva table es

$$\begin{array}{ccccc|c} x_1 & x_4 & x_2 & x_3 & x_5 & \\ 1 & 0 & 0 & 1/3 & -2/3 & 6 \\ 0 & 1 & 0 & -1/3 & 2/3 & 4 \\ 0 & 0 & 1 & 0 & 1 & 6 \\ \hline 0 & 0 & 0 & 1 & 3 & 36 \end{array}$$

Dado que en esta tabla se cumple el criterio de parada (todas las componentes no nulas de c son no negativas), la solución óptima se encuentra en la columna b. El costo óptimo (con el signo cambiado) se encuentra en d, -36*. Es importante determinar los componentes asociados con los valores de b. En esta ultima tabla la matriz B es formada por tres columnas* x_1*,* x_4 *y* x_2*, y las componentes de b correspondenderan (en ese orden) a esas variables. A las variables ausentes en B se les asigna el valor de* 0*. Por lo tanto la solución optima es* $\begin{pmatrix} 2 & 6 & 0 & 2 & 0 \end{pmatrix}$ *con costo óptimo* -36*. En la practica, los calculos se hacen transformadon las tablas sin mayor comentario.*

Resolvemos otro ejemplo incluyendo la discusión en la inicialización por el segundo metodo que hemos indicado antes.

Ejemplo 2.17 *El problema es*

$$\textit{Mínimizar} \quad 3x_1 + x_2 + 9x_3 + x_4$$

sujeto a

$$\begin{aligned} x_1 + 2x_3 + x_4 &= 4 \\ x_2 + x_3 - x_4 &= 2 \\ x_i &\geq 0. \end{aligned}$$

De acuerdo a nuestra discución de como plantear un nuevo problema de optimización equivalente para el cual la inicialización es trivial, consideramos el nuevo PPL modificado

$$\textit{Mínimizar} \quad 3x_1 + x_2 + 9x_3 + x_4 + dx_5 + dx_6$$

sujeto a

$$\begin{aligned} x_1 + 2x_3 + x_4 + x_5 &= 4 \\ x_2 + x_3 - x_4 + x_6 &= 2 \\ x_i &\geq 0. \end{aligned}$$

donde se asume que d es un parametro muy grande. La tabla inicial para este problema es

$$\begin{array}{cccccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & \\ 1 & 0 & 2 & 1 & 1 & 0 & 4 \\ 0 & 1 & 1 & -1 & 0 & 1 & 2 \\ & & & & & & \\ 3 & 1 & 9 & 1 & d & d & 0 \end{array}$$

Notamos que una inicialización admisible es la matriz identidad correspondiente a la quinta y sexta columna. La segunda tabla es

$$\begin{array}{cccccccc} x_5 & x_6 & x_1 & x_2 & x_3 & x_4 & \\ 1 & 0 & 1 & 0 & 2 & 1 & 4 \\ 0 & 1 & 0 & 1 & 1 & -1 & 2 \\ & & & & & & \\ 0 & 0 & 3-d & 1-d & 9-3d & 1 & -6d \end{array}$$

Si d es un numero muy grande positivo, el coeficiente mas negativo de c sera $9 - 3d$*, asi que escogemos a* x_3 *como la variable entrante, y entonces* x_1 *como la variable saliente (en este caso particular las dos razones son iguales, y podiamos igualmente escoger a* x_6 *como variable saliente). Despues de reorganizar columnas y hacer algunos calculos tenemos.*

$$\begin{array}{cccccccc} x_3 & x_6 & x_1 & x_2 & x_5 & x_4 & \\ 1 & 0 & 1/2 & 0 & 1/2 & 1/2 & 2 \\ 0 & 1 & -1/2 & 1 & -1/2 & -3/2 & 0 \\ & & & & & & \\ 0 & 0 & d-3/2 & 1-d & 3(d-3)/2 & (3d-7)/2 & -18 \end{array}$$

De nuevo teniendo en cuenta que d es un numero muy grande y positivo, escogeriamos x_2 *como la variable entrante y* x_6 *como nuestra variable saliente. Despues de hacer los calculos la tabla es*

$$\begin{array}{cccccccc} x_3 & x_2 & x_1 & x_6 & x_5 & x_4 & \\ 1 & 0 & 1/2 & 0 & 1/2 & 1/2 & 2 \\ 0 & 1 & -1/2 & 1 & -1/2 & -3/2 & 0 \\ & & & & & & \\ 0 & 0 & -1 & d-1 & d-4 & -2 & -18 \end{array}$$

Nuestra siguiente variable entrante es x_4*, y* x_3 *es nuestra variable saliente. Asi*

obtenemos

$$\begin{array}{cccccccc} x_4 & x_2 & x_1 & x_6 & x_5 & x_3 & \\ 1 & 0 & 1 & 0 & 1 & 2 & 4 \\ 0 & 1 & 1 & 1 & 1 & 3 & 6 \\ \\ 0 & 0 & 1 & d-1 & d-2 & 4 & 10 \end{array}$$

Dado que todos los componentes en la ultima fila son positivos hemos llegado a una solución óptima, que esta dada por la ultima columna 4, 6 *para* x_4 *y* x_2*, respectivamente, y a las demas variables debe asignarseles el valor* 0*. El costo óptimo es 10 y la solución óptima* $\begin{pmatrix} 0 & 6 & 0 & 4 & 0 & 0 \end{pmatrix}$*. Note que esta solución óptima tiene componentes nulas para las variables auxiliares* x_5 *y* x_6*. La solución óptima para el problema original será* $\begin{pmatrix} 0 & 6 & 0 & 4 \end{pmatrix}$ *con costo óptimo* 10*.*

Alternativamente podemos asignar un valor numérico muy alto a d (mucho mas grande que aquellos participando en el problema por ejemplo $d = 100$*) y resolver el problema. Si la solución final da valores nulos para* x_5 *y* x_6*, tenemos nuestra solución óptima. Si no, se debe resolver de nuevo el problema con un valor mayor de d.*

Finalmente, queremos enfatizar de manera mas explicita como el pasar de la solución optima del primario a la solución optima del dual se puede hacer de una manera eficiente. De hecho, esto se trato cuando vimos dualidad, pero nos gustaria enfatizar en esto aqui. Asuma que el primario es

$$\text{Mínimizar} \quad cx \quad \text{bajo} \quad Ax = b, \quad x \geq 0,$$

con dual

$$\text{Máximizar} \quad yb \quad \text{bajo} \quad yA \leq c.$$

Si

$$x = \begin{pmatrix} x_B & 0 \end{pmatrix}, \quad x_B = B^{-1}b$$

es la solución óptima del primario, la solución óptima del dual sera

$$y = c_B B^{-1},$$

donde c_B incorpora las componentes de c correspondientes a las columnas seleccionadas en B. En la practica, es cuestion de resolver el sistema lineal

$$c_B = yB$$

donde B y c_B incluyen las columnas para la solución optima. Es mas, estas columnas estan asociadas con las desigualdades que deben convertirse en igualdades para encontrar la solución optima del dual. Un ejemplo aclarara esto ultimo.

Ejemplo 2.18 *Queremos mínimizar la función* $18y_1 + 4y_2 + 6y_3$ *bajo las resticciones*

$$3y_1 + y_2 \leq -3, \quad 2y_1 + y_3 \leq -5, \quad y \leq 0$$

En forma maticial estas restricciones se pueden escribir como

$$\begin{pmatrix} 2 & 1 & 0 \\ 2 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -3 \\ -5 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Hemos usado la variable y para sugerir que este problema puede ser entendido directamente como el dual de cierto PPL primario. Es cierto que podemos resolverlo transformandolo a la forma estándar y aplicando el metodo simplex. Pero este proceso requiere mas esfuerzo que si lo tratamos como un problema dual. De hecho el primario asociado es

$$\textit{Mínimizar} \quad -3x_1 - 5x_2$$

sujeto a

$$\begin{aligned} 3x_1 + 2x_2 + x_3 = 18, \quad x_1 + x_4 = 4 \\ x_2 + x_5 = 6, \quad x \geq 0. \end{aligned}$$

Ya se ha resuelto este problema, Su solución optima es el vector

$$\begin{pmatrix} 2 & 6 & 0 & 2 & 0 \end{pmatrix}$$

Las componentes no nulas de este vector indican que la matrix B incluye la primera, segunda y cuarta columnas de A. Si hubieramos tenido dos (o menos) componentes no nulas, la tercera columna se podria escoger de manera arbitraria mientras la matriz resultante sea no singular. Esta información es suficiente para resolver el dual, dado que su solución optima se puede encontrar resolviendo el sistema lineal

$$\begin{pmatrix} 3 & 1 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -3 \\ -5 \\ 0 \end{pmatrix}$$

obtenidas transformando en igualdades la primera, segunda y cuarta desigualdades del dual. La solución óptima es entonces $\begin{pmatrix} -1 & 0 & 3 \end{pmatrix}$ *y el costo óptimo es* -36

2.5. Programación entera

Las aplicaciones a la ingeniería de la programación lineal requieren a menudo que las variables tomen valores enteros en vez de reales. En algunos casos, ignorar esta restricción dara una aproximación razonable. En otros, es crucial poner atención a esta restricción. En estos casos, ademas de las tipicas restricciones lineales

$$Ax = b, \quad x \geq 0$$

debemos forzar a algunas (o todas) las variables tomar valores enteros (no negativos) $x_i \in \mathbb{Z}$. Esta restricción nueva estara dada por la naturaleza del problema en la que estamos interesados. Nos enfrentamos por lo tanto al PPL que identificaremos como $(\tilde{P})$,

$$\begin{aligned} \text{Mínimizar} \quad cx \quad \text{bajo} \quad Ax = b, \quad x \geq 0, \\ x_i \in \mathbb{Z}, \quad i \in I \subset N = \{1, 2, \ldots, n\} \end{aligned}$$

donde el subvonjunto de indices I es conocido. De forma razonable, nos preocuparemos primero acerca del PPL esencial sin la restricción entera.

$$\text{Mínimizar} \quad cx \text{bajo} \quad Ax = b, \quad x \geq 0.$$

Asumamos por el momento que $x^{(0)}$, la solución óptima para (P), satisface el requerimiento entero $x_i^{(0)} \in \mathbb{Z}$. En este caso es evidente que hemos encontrado la (o una) solución óptima para $(\tilde{P})$. Sin embargo es muy probable que tengamos tanta suerte, y que tal solución óptima no verificara la condición entera. ¿Cómo proceder en tal situación? La estrategia a seguir es llamada "metodo de dividir y acotar", y consiste en generar una secuencia de subproblemas, resolverlos y analizar y comparar las distintas soluciones hasta que lleguemos a una solución óptima y factible a nuestro problema original.

La idea basica detras de descomponer un problema en dos problemas disyuntos ("dividir") es la siguiente. Asuma por recursión que tenemos un PPL como resultado de los pasos previos y no hemos encontrado un vector factible para nuestro problema inicial. Encontramos su solución óptima $x^{(0)}$. Obviamente si el problema es insatisfactible es descartado. Pueden ocurrir dos situaciones.

1. Si $x^{(0)}$ satisface las restricciones enteras se convierte en nuestra solución provisional y descartamos el subproblema.

2. Si $x^{(0)}$ no satisface todas las restricciones enteras, escogemos una variable

   $$x_i^{(0)} \in (k, k+1), \quad k \in \mathbb{Z}, \quad i \in I,$$

   y aadimos a la colección de subproblemas a ser analizados los dos problemas disyuntos ("dividir") obteniendo aadiendo a las restricciones del problema las restricciones $x_i \leq k$ en un caso, y $x_i \geq k+1$ en el otro. Note como los conjuntos factibles para estos dos subproblemas son disyuntos y su union es el PPL completo del que vienen. Vea la figura 2.6.

Una vez hemos encontrado una solución provisional óptima factible x^*, y tenemos que analizar un subproblema previamente generado por el proceso de dividir, la discusión será la siguiente.

1. Si

   $$cx^* \leq cx^{(0)}$$

   descartamos este PPL, dado que no puede mejorar la solución óptima que encontramos y elegimos otro subproblema.

2. Si
$$cx^* > cx^{(0)}$$
y $x^{(0)}$ satisface el requisito entero, cambie la solución óptima provisional a $x^{(0)}$, descarte el problema correspondiente y ana'lice otro problema

3. Si
$$cx^* > cx^{(0)}$$
y $x^{(0)}$ no satisface la restricción entera, proceda a dividir este problema como se indico anteriormente. De esta manera estamos asegurando que la solución óptima se encontrara por estre proceso exhaustivo. De nuevo cualquier subproblema que son no factibles debido a la falta de vectores factibles son eliminados.

   Despues que todos los subproblemas se han analizado, la solución provisional óptima se convierte en la solución óptima de nuestro problema inicial. Este es siempre un proceso finito.

Ejemplo 2.19 *Queremos resolver el problema* $(\tilde{P})$

$$\textit{Mínimizar} \quad 3x_2 + 2x_3$$

bajo

$$\begin{aligned} 2x_1 + 2x_2 - 4x_3 = 5 \quad 4x_2 + 2x_3 \leq 3, \\ x_i \geq 0. \qquad x_1, x_3 \in \mathbb{Z} \end{aligned}$$

El PPL de fondo es

$$\begin{aligned} \textit{Mínimizar} \quad 3x_2 + 2x_3 \quad \textit{bajo} \\ 2x_1 + 2x_2 - 4x_3 = 5, \quad 4x_2 + 2x_3 \leq 3, \quad x_i \geq 0. \end{aligned}$$

Cuya solución óptima es $\begin{pmatrix} 5/2 & 0 & 0 \end{pmatrix}$. *Dado que* x_i *no es un entero, debemos "dividir" el problema. Hasta ahora no tenemos una solución factible provisional. Los dos subproblemas son*

$$\begin{aligned} \textit{Mínimizar} \quad 3x_2 + 2x_3 \quad \textit{bajo} \\ 2x_1 + 2x_2 - 4x_3 = 5, \quad 4x_2 + 2x_3 \leq 3, \\ x_1 \leq 2, \quad x_i \geq 0. \end{aligned}$$

y

$$\begin{aligned} \textit{Mínimizar} \quad 3x_2 + 2x_3 \quad \textit{bajo} \\ 2x_1 + 2x_2 - 4x_3 = 5, \quad 4x_2 + 2x_3 \leq 3, \\ x_1 \geq 3, \quad x_i \geq 0. \end{aligned}$$

Sus respectivas soluciones óptimas son

$$\begin{pmatrix} 2 & \frac{1}{2} & 0 \end{pmatrix}, \quad coste = \frac{3}{2}$$
$$\begin{pmatrix} 3 & 0 & \frac{1}{4} \end{pmatrix}, \quad coste = \frac{1}{2}$$

El primero de estos respecta la condicion entera, y por lo tanto es tomado como nuestra solución óptima provisional. Dado que el coste del segundo es menor que el coste de sa solución provisional debemos considerar este subproblema, puesto que podría contener una mejor solución. Por otro lado, dado que la segunda solución no respecta la condición entera debemos dividir este problema y obtenemos.

$$\begin{aligned} &\textit{Mínimizar} \quad 3x_2 + 2x_3 \quad \textit{bajo} \\ 2x_1 + 2x_2 - 4x_3 = 5, \quad &4x_2 + 2x_3 \leq 3, \\ x_1 \geq 3, \quad x_3 \leq 0, \quad &x_i \geq 0. \end{aligned}$$

y

$$\begin{aligned} &\textit{Mínimizar} \quad 3x_2 + 2x_3 \quad \textit{bajo} \\ 2x_1 + 2x_2 - 4x_3 = 5, \quad &4x_2 + 2x_3 \leq 3, \\ x_1 \geq 3, \quad x_3 \geq 1, \quad &x_i \geq 0. \end{aligned}$$

El primero es insatisfactible dado que $x_3 = 0$, y de la primera restricción tenemos $2x_1 + 2x_2 = 5$. Esto junto con $x_1 \geq 3$ y $x_2 \geq 0$ es imposible. Este subproblema debe ser descartado. La solución optima para el segundo es $\begin{pmatrix} 9/2 & 0 & 1 \end{pmatrix}$ con costo 2. Desde que este costo es mayor que el de la solución provisional, eliminamos este problema sin cambiar la solución provisional. Dado que no hay mas subproblemas que solucionar, la solución provisional es la solución optima para el problema inicial.

$$\begin{pmatrix} 2 & \frac{1}{2} & 0 \end{pmatrix}, \quad coste = \frac{3}{2}$$

Note como difiere de la solución optima sin la restricción entera.

En la practica, no es facil decidir en que variable debe ser dividida de tal modo que el proceso resulta tan corto y eficiente como sea posible. No hay reglas fijas para determinar la escogencia mas eficiente y cualquier situación. Cualquier información disponible a priori del problema puede decidir el camino a escoger entre las alternativas disponibles.

2.6. Ejercicios

1. Dibuje en el plano la región deteminada por las desigualdades

$$x_2 \geq 0, \quad 0 \leq x_1 \leq 3, \quad -x_1 + x_2 \leq 1, \quad x_1 + x_2 \leq 4.$$

En el (los) punto(s) donde las siguientes funiones tienen sus valores mínimos y máximos.

$$2x_1 + x_2, \quad x_1 + x_2, \quad x_1 + 2x_2.$$

2. Resuelva graficamente los siguientes dos problemas:

$$\text{Máximizar} \quad 2x_1 + 6x_2$$

sujeto a

$$-x_1 + x_2 \leq 1, \quad 2x_1 + x_2 \leq 2, \quad x_1 \leq 0, \quad x_2 \geq 0.$$

$$\text{Mínimizar} \quad -3x_1 + 2x_2$$

sujeto a

$$x_1 + x_2 \leq 5, \quad 0 \leq x_1 \leq 4, \quad 1 \leq x_2 \leq 6.$$

3. Determine los valores del parametro d tal que el conjunto factible determinado por

$$x_1 + x_2 + x_3 \leq d, \quad x_1 + x_2 - x_3 = 1, \quad 2x_3 \geq d,$$

es vacio.

4. Determine los vectores donde la función lineal $2x_1 + 3x_2 + x_3$ toma su máximo bajo las restricciones

$$\begin{aligned} x_1 \geq 0, \quad x_2 \geq 0, \quad x_2 \geq 0, \\ x_1 + x_2 + 2x_3 \geq 200, \\ 3x_1 + 2x_2 + 2x_3 \leq 300. \end{aligned}$$

5. El valor máximo de la función $3x_1 + 2x_2 - 2x_3$ se busca con respecto a las restricciones

$$\begin{aligned} 4x_1 + 2x_2 + 2x_3 \leq 20, \\ 2x_1 + 2x_2 + 4x_3 \geq 6, \\ x_1 \geq 0, \quad x_2 \geq 0, \end{aligned}$$

pero el signo de x_3 no esta restringido. Encuentre la(s) solución(es) óptima(s).

6. Determine el valor máximo de $18x_1 + 4x_2 + 6x_3$ bajo las restricciones

$$3x_1 + x_2 \leq -3, \quad 2x_1 + x_3 \leq -5, \quad x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0,$$

revisando el problema dual.

7. Considere el siguiente problema primario:

$$\text{Máximizar} \quad 1{,}1x_1 + 1{,}2x_2 + x_3$$

sujeto a

$$2x_1 + 2x_2 + 2x_3 \leq 10, \quad x_1 + 3x_2 + x_3 \leq 10, \quad 4x_1 + x_2 + x_3 \leq 10$$
$$3x_1 + x_2 + 3x_3 \leq 10, \quad x_1 + 2x_2 + 3x_3 \leq 10, \quad 3x_1 + 2x_2 + x_3 \leq 10,$$
$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0.$$

Formule el dual y resuelvalo.

8. Encuentre explicitamente la solución óptima del PPL

$$\text{Mínimizar} \quad x_1 + x_2 + x_3$$

sujeto a

$$x_1 + 2x_2 + 3x_3 = b_1, \quad x_1 - x_2 - x_3 = b_2,$$

en terminos de b_1 y b_2. Encuentre la solución óptima del problema dual y revise su relación con el gradiente del valor óptimo del primario con respecto a b_1 y b_2.

9. Resuelva el problema del tejado en el capítulo 1, calcule las cantidades de cada modelo que deben ser enviadas para máximizar los beneficios.

10. Resuelva el ejercicio del sistema de resortes del capítulo 1 para el siguiente conjunto de datos.

    a) ubicacion de los nodos fijos. $(1, 0)$, $(0, 1)$, $(-1, 0)$, $(0, -1)$.
    b) $k = 1$
    c) $F = (1, 1)$.

11. Resuelva el problema del transporte del capítulo 1 con el siguiente conjunto de datos

    a) $n = 3$, $m = 2$.
    b) $u_1 = 2$, $u_2 = 2$, $u_3 = 2$, $v_1 = 5$, $v_2 = 2$.
    c) $c_{11} = 2$, $c_{12} = 1$, $c_{21} = 3$, $c_{22} = 1$, $c_{31} = 2$, $c_{32} = 3$.

12. Trate de describir la mejor solución al problema de inversión en el capítulo 1.

13. Resuelva el problema del andiamaje propuesto en el capítulo 1, donde las cargas x_1 y x_2 son aplicadas exactamente en los puntos medios de las barras CD y EF, respectivamente.

14. Aunque hay muchos paquetes informaticos para resolver PPL de dimensión grande, no es muy dificil disear un programa para implementar el metodo simplex. Hagalo en algu lenguaje (C, Fortran, Maple, Mathematica, Matlab, etc) y uselo para resolver los siguientes problemas.

a) Máximizar $x_1 + x_2 - x_6$ sujeto a

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0, \quad x_5 \geq 0, \quad x_6 \geq 0,$$
$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1.$$

b) Máximizar $x_1 - x_2 + x_3 - x_4 + x_5 - x_6$ bajo las restricciones

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0, \quad x_5 \geq 0, \quad x_6 \geq 0,$$
$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \geq -1.$$

c) Máximizar $x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 + 6x_6$ bajo las restricciones

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0, \quad x_4 \geq 0, \quad x_5 \geq 0, \quad x_6 \geq 0,$$
$$6x_1 + 5x_2 + 4x_3 + 3x_4 + 2x_5 + x_6 \leq 1, \quad 6x_1 + x_2 + 5x_3 + 2x_4 + 4x_5 + 3x_6 \leq 1$$

d) Máximizar $x_1 - x_2 + x_3 - x_4 + x_5 - x_6 + x_7 - x_8 + x_9 - x_{10}$ bajo las restricciones

$$-1 \leq x_1 + x_2 \leq 1, \quad -1 \leq x_1 + x_2 + x_3 \leq 1,$$
$$-1 \leq x_2 + x_3 + x_4 \leq 1, \quad -1 \leq x_3 + x_4 + x_5 \leq 1,$$
$$-1 \leq x_4 + x_5 + x_6 \leq 1, \quad -1 \leq x_5 + x_6 + x_7 \leq 1,$$
$$-1 \leq x_6 + x_7 + x_8 \leq 1, \quad -1 \leq x_7 + x_8 + x_9 \leq 1,$$
$$-1 \leq x_8 + x_9 + x_{10} \leq 1, \quad -1 \leq x_9 + x_{10} \leq 1.$$

15. Algunas funciónes no linales se pueden tratar en el contexto de los PPL. Trate de formular y resolver el siguiente problema.

$$\text{Mínimizar} \quad |x_1| - x_2$$

sujeto a

$$x_1 + |x_2| \leq 1, \quad 2|x_1| - |x_2| \leq 2.$$

(Pista: La función $|\cdot|$ puede ser modelada de forma lineal descomponiendola en la suma de dos variables independientes no negativas, de la misma manera que una variable que no esta restringida en signo es la differencia de dos tales variables).

16. Considere el PPL simple.

$$\text{Máximizar} \quad x_1 + 2x_2$$

sujeto a $x_1 + x_2 \leq 1$, $0 \leq x_1 \leq 1$, $0 \leq x_2$. Muestre que el metodo simplex entra a un proceso cíclico escogiendo como inicialización la matriz correspondiente a las variables x_1, x_2. Note como la desigualdad $x_1 \leq 1$ es redundante con $x_1 + x_2 \leq 1$, $0 \leq x_1, x_2$. Pruebe si eliminando esa desigualdad, el metodo simplex evita el proceso cíclico.

Capítulo 3

Programación no lineal

3.1. Problema modelo

El problema que trataremos en este capítulo tiene una estructura similar a la de un PPL. Nos gustaria aprender como

$$\text{Mínimizar} \quad C(x) \quad \text{bajo} \quad A(x) \leq 0.$$

En la situación de un PPL ambos el funcional de costo C y las funciones que determinan la admisibilidad A eran lineales. Si cualquiera C o alguno de los componentes de A fueran no lineales, el problema anterior de dice que esun problema de programación no lineal (PPNL). Como nuestros lectores pueden facilmente inferir, estos problemas son considerablemente mas complejos que sus contrapartes lineales. Suponemos que todas las funciones son suaves a no ser que se diga lo contrario.

A pesar que hemos escrito las restricciones en la forma de desigualdades, estas se puden expresar como igualdades y/o desigualdades como ha sido mencionado en el capítulo anterior, donde enfatizamos que multiplicando por -1 cambia la dirección de una desigualdad, y que esa igualdad es equivalente a dos desigualdades. Dado que las restricciones en forma de igualdades y deigualdades juegan un papel importante el PPNL, ellas tipicamente son distinguidas con diferentes nombres, a traves de este capítulo nos apegaremos a la siguiente forma general de un PPNL.

Definición 3.1 *Forma general de un PPNL La forma estándar de un PPNL es*

$$\textit{Mínimizar} \quad f(x) \quad \textit{sujeto a} \quad g(x) \leq 0, \quad h(x) = 0.$$

donde $x \in \mathbb{R}^n$.

Una primera primera pregunta esta relacionada con la existencia de soluciones óptimas para este problema. Ya sabemos que incluso un PPL puede no tener soluciones óptimas. Esto tambien es cierto para un PPNL. Un resultado típico

asegura la existencia de soluciones óptimas esta basado en la continuidad de las funciones involucradas en tal problema.

Teorema 3.2 *Suponga que f, g y h son funciones continuas y una de las siguientes dos situaciones ocurre.*

1. *El conjunto de vectores factibles $g(x) \leq 0$, $h(x) \leq 0$ es un conjunto acotado en $\mathbb{R}^n$.*

2. *El conjunto de vectores factibles no esta acotado pero*

$$\lim_{|x|\to\infty, g(x)\leq 0, h(x)=0} f(x) = +\infty.$$

Entonces el problema de minimización asociado admite al menos una solución.

Hay mas situaciones en las cuales un PPNL admite una solución óptima, pero dilucidar eso es parte de la meta de este capítulo (Sección 5).

En varias situaciones practicas, el teorema anterior es suficiente para asegurar la existencia de una solución óptima. El tema principal en esta capítulo es como encontrarla.

Para entender mejor como encontrar o aproximar soluciones óptimas procederemos en dos pasos. Primero, tratartemos el caso en que las restricciones vienen en la forma de igualdades.

$$\text{Mínimizar} \quad f(x) \quad \text{sujeto a} \quad h(x) = 0.$$

Entones examinaremos el caso general aplicando apropiadamente la situación de las restricciones de igualdad. El asunto principal que nos gustaria entender es, ¿Qué es especial acerca de las soluciones óptimas de un PPNL?, ¿Qué deben satisfacer para ser elegibles como solucion óptima para un problema de optimización particular? Esta es la pregunta acerca de las condiciones necesarias para la optimalidad, y nos llevara a las condiciones de Karush-Kuhn Tucker (KKT). Investigaremos varios ejemplos explicitos. A continuación nos interesatemos en situaciones en las cuales estas condiciones necesarias de optimalidad son de hecho suficientes para detectar soluciones óptimas (globales) a un PPNL. Esto abrirá el problema de entender la convexidad, y por que es una propiedad deseable al minimizar un funcional de costo bajo un conjunto de restricciones. Terminaremos el capítulo con una breve descución acerca de la dualidad para un PPNL.

En el tratamiento de un PPNL es importante hacer una distinción entre los minimos locales y globales

Definición 3.3 *Un vector $x^{(0)} \in \mathbb{R}^n$ es un mínimo local de f sujeto a $g(x) \leq 0$, $h(x) = 0$ si.*

$$g(x^{(0)}) \leq 0, \quad h(x^{(0)}) = 0.$$

y

$$f(x^{(0)}) \leq f(x)$$

Para todo x tal que

$$g(x) \leq 0, \quad h(x) = 0, \quad |x - x^{(0)}| \leq \epsilon$$

para algun $\epsilon > 0$.

Un vector $x^{(0)}$ es un mínimo global de f sujeto a $g(x) \leq 0$, $h(x) = 0$ si

$$g(x^{(0)}) \leq 0, \quad h(x^{(0)}) = 0.$$

y

$$f(x^{(0)}) \leq f(x)$$

Para todo x tal que

$$g(x) \leq 0, \quad h(x) = 0.$$

Note la diferencia entre estos dos conceptos.

3.2. Multiplicadores de Lagrange

En esta sección trataremos de derivar las condiciones que debe satrisfacer un vector para que sea posible que sea una solución óptima para el problema

$$\text{Mínimizar} \quad f(x) \quad \text{bajo} \quad h(x) = 0.$$

Puede ser posible que algunos de nuestros lectores sepan de antemano como escribir las condiciones de optimalidad para el problema anterior. i.e. aquellas ecuaciones en terminos de f y h que deben satisfacer las soluciones óptimas. Esto se ensea en ocasiones en cursos de cálculo avanzado. Se necesitan introducir los multiplicadores de Lagrange los cuales son parametros asociados a las restricciones, uno por cada constante individual. Si λ es tal vector de multiplicadores para el sistema de ecuaciones.

$$\nabla f(x) + \lambda \nabla h(x) = 0, \quad h(x) = 0. \tag{3.1}$$

donde los pares (x, λ) de puntos y multiplicadores son las incognitas. Note que tenemos tantas ecuaciones como incognitas $n+m$, si $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^m$ y tenemos m restricciones, de tal manera que $h : \mathbb{R}^n \to \mathbb{R}^m$. Es importante enfatizar que no todas las soluciones de (3.1) seran soluciones óptimas para nuestro problema. Lo que es cierto es quen nuestras soluciones óptimas estan entre las soluciones de (3.1). Otras soluciones de este systema pueden corresponder a máximos, mínimos y máximos locales, puntos de silla, etc.

Teorema 3.4 *Toda solución óptima de*

$$\textit{Mínimizar} \quad f(x) \quad \textit{bajo} \quad h(x) = 0$$

debe ser una solución del sistema de condiciones necesarias de optimalidad

$$\nabla f(x) + \lambda \nabla h(x) = 0, \quad h(x) = 0.$$

Antes de dar alguna justificación para las condiciones de optimalidad a aquellos lectores interesados, vamos a mirar varios ejemplos y ver como se pueden usar para encontrar soluciones óptimas.

Ejemplo 3.5 *Nos gustarua encontrar los valores extremos (máximos y mínimos) de la función*

$$f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$$

sobre la esfera $x_1^2 + x_2^2 + x_3^2 = 4$. *En este caso* $n = 3$, $m = 1$, *y*

$$h(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 4$$

Las condiciones de optimalidad (3.1) pueden ser escritas

$$\begin{aligned} 3x_1^2 + 2\lambda x_1 &= 0, \\ 3x_2^2 + 2\lambda x_2 &= 0, \\ 3x_3^2 + 2\lambda x_3 &= 0, \\ x_1^2 + x_2^2 + +x_3^2 &= 4. \end{aligned}$$

Escribir estas ecuaciones no debe representar ninguna dificultad. Aunque encontrar sus soluciones puede requerir, cierta habilidad computacional. Dado que tenemos que enfrentarnos con un sistema de ecuaciones, no hay manera de saber de antemano cuantos vectores estamos buscando, asi que en la manipulación de las ecuaciones tenemos que asegurar que no se pierde ninguna solución, dado que en particular la solución óptima que estamos buscando puede ser precisamente aquella no encontrada. En nuestra situación particular, factorizando las primeras tres ecuaciones, obtenemos

$$\begin{aligned} (3^2 + 2\lambda)x_1 &= 0, \\ (3^2 + 2\lambda)x_2 &= 0, \\ (3^2 + 2\lambda)x_3 &= 0, \\ x_1^2 + x_2^2 + +x_3^2 &= 4. \end{aligned}$$

Las primeras tres ecuaciones tienen son productos, y de esta manera tendremos ocho posibilidades dependiendo en los factores que se anulan. Pero, dada la simetria de las ecuaciones con respecto a las tres variables independientes, nos basta con considerar cuatro casos

1. $x_1 = x_2 = x_3 = 0$*: esta situación es inconsistente con la restricción.*
2. $x_1 = x_2 = 0$, $x_3 \neq 0$*: teniendo en cuenta la restricción obtenemos* $x_3 = \pm 2$, *y la tercera ecuación puede usarse para determinar el valor del multiplicador (no nos interesa encontrarlo por el momento).*
3. $x_1 = 0$, $x_2, x_3 \neq 0$*: de la segunda y tercera ecuaciones concluimos que* $x_2 = x_3$ *y usando esto en la restricción,* $x_2 = x_3 = \pm\sqrt{2}$.
4. $x_1, x_2, x_3 \neq 0$*: las primeras tres ecuaciones nos llevan a* $x_1 = x_2 = x_3$, *y la restricción nos aegura que el valor comun es* $\pm 2/\sqrt{3}$.

Teniendo en cuenta la simetria tenemos los siguientes candidatos para puntos máximos y mínimos.

$$(\pm 2,0,0),\quad (0,\pm 2,0),\quad (0,0,\pm 2),(0,\sqrt{2},\sqrt{2})\quad (0,-\sqrt{2},-\sqrt{2})\quad (\sqrt{2},0,\sqrt{2})$$
$$(-\sqrt{2},0,-\sqrt{2})\quad (\sqrt{2},\sqrt{2},0)\quad (-\sqrt{2},-\sqrt{2},0),\quad (2/\sqrt{3},2/\sqrt{3},2/\sqrt{3}),$$
$$(-2/\sqrt{3},-2/\sqrt{3},-2/\sqrt{3}).$$

Ademas, la observación importante que la esfera es una superficie acotada en el espacio nos permite saber que de hecho la función continua f debe alcanzar sus dos valores extremos en alguna parte. Por lo tanto los puntos donde se alcanzan el máximo y el mínimo estan contenidos en la lista anterior. Simplemente calculando f en esos puntos y comparando sus valores, encontramos que el máximo es 8 y se alcanza en $(2,0,0)$, $(0,2,0)$, $(0,0,2)$ mientras que el mínimo es -8 y corresponde a $(-2,0,0)$, $(0,-2,0)$, $(0,0,-2)$.

Ejemplo 3.6 *Supongamos que estamos interesados en conocer los valores extremos (máximos y mínimos) de la misma función f no sobre todos los puntos de la esfera, sino sobre aquellos que estan al mismo tiempo en el plano $x_1+x_2+x_3=1$, esto es, nos interesa encontrar los puntos extremos de la función*

$$f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2$$

sobre el conjunto de puntos que satisface

$$x_1^2+x_2^2+x_3^2=4,\quad x_1+x_2+x_3=1.$$

Esta vez las ecuaciones de optimalidad son

$$\begin{aligned}3x_1^2+2\lambda_1x_1+\lambda_2&=0,\\ 3x_2^2+2\lambda_1x_2+\lambda_2&=0,\\ 3x_3^2+2\lambda_1x_3+\lambda_2&=0,\\ x_1^2+x_2^2+x_3^2&=4,\\ x_1+x_2+x_3&=1.\end{aligned}$$

un sistema de cinco ecuaciones con cinco incognitas $x_1,x_2,x_3,\lambda_1,\lambda_2$. Dado que no estamos particularmente interesados en encontrar los valores de los multiplicadores, y que las tres primeras ecuaciones implican que ek gradiente de f debe ser una combinación linal de los gradientes de las dos restricciones, podemos reescribir el sistema anterior como

$$\begin{aligned}\begin{vmatrix}3x_1^2 & 2x_1 & 1\\ 3x_2^2 & 2x_2 & 1\\ 3x_3^2 & 2x_3 & 1\end{vmatrix}&=0,\\ x_1^2+x_2^2+x_3^2&=2,\\ x_1+x_2+x_3&=1.\end{aligned}$$

Despues de factorizar el 3 *en la primera columna y el* 2 *en la segunda, obtenemos un determinante de Vandermonde, cuya expresión es bien conocida*

$$\begin{aligned}(x_1 - x_2)(x_1 - x_3)(x_2 - x_3) &= 0,\\ x_1^2 + x_2^2 + x_3^2 &= 4,\\ x_1 + x_2 + x_3 &= 1.\end{aligned}$$

Es elemental encontrar las soluciones de este sistema. Hay tres posibilidades, que debido a la simetria del problema se reducen escencialmente a una:

$$x_1 = x_2, \quad x_3 = 1 - 2x_1, \quad 2x_1 + (1 - 2x_1)^2 = 4.$$

Las otras soluciones son obtenidas por permutaciones de las variables. Las soluciones explicitas son

$$\begin{aligned}&(\frac{1}{3} + \frac{\sqrt{22}}{6}, \frac{1}{3} + \frac{\sqrt{22}}{6}, \frac{1}{3} - \frac{\sqrt{22}}{6}),\\ &(\frac{1}{3} - \frac{\sqrt{22}}{6}, \frac{1}{3} - \frac{\sqrt{22}}{6}, \frac{1}{3} + \frac{\sqrt{22}}{6}),\end{aligned}$$

y aquellas obtenidas por permutaciones de estas. Es directo comprobar que el primer conjunto de soluciones corresponde al valor máximo, y los otros al minimo. Note como estas soluciones difieren de aquellas del ejemplo 3.5.

¿Cómo podemos entender de donde vienen estos multiplicadores? Una forma simple de entender esto es considerar las curvas parametrizadas

$$\tau(-\delta, \delta) \to \mathbb{R}^n, \quad \delta > 0,$$

cuya imagen $\tau(-\delta, \delta)$ esta enteramente contenido en el conjunto factible de nuestro problema de optimización, esto es

$$h(\tau(t)) = 0, \quad \text{para todo} \quad t \in (-\delta, \delta).$$

Si suponemos que $x_0 \in \mathbb{R}^n$ es mínimo o máximo local, incluso un punto de silla con respecto a los vectores del conjunto factible, y suponga que τ pasa por x_0, para $t = 0$, $\tau(0) = x_0$, entonces la composición $f(\tau(t))$ debe de la misma manera tener un mínimo, máximo o punto de silla en $t = 0$. Lo que caracteriza a cualquiera de estas situaciones es que la derivada se debe anular. Por la regla de la cadena,

$$0 = \left.\frac{df(\tau(t))}{dt}\right|_{t=0} = \nabla f(\tau(0))\tau'(0) = \nabla f(x_0)\tau'(0)$$

Por otro lado, dado que $h(\tau(t)) = 0$ para todo t, debemos tener de la misma manera

$$0 = \nabla h(x_0)\tau'(0).$$

Dado que el vector $\tau'(0)$ es arbitrario, en cuanto respecta a estas dos igualdades, comcluimos que estas se pueden dar simultaneamente si y solo si $\nabla f(x_0)$ pertenece al generado de $\nabla h(x_0)$. Esta dependencia lineal de lugar a los multiplicadores. Vease la figura 3.1.

Es importante sealar que el metodo de los multiplicadores puede fallar en hallar los puntos extremos cuando algunas de las restricciones $h_i(x) = 0$ representan una superficie (o hipersuperficie) que no es regularm en el sentido que su vector gradiente ∇h_i, o cuando la intersección de los conjuntos $h_i(x) = 0$ es de alguna manera no regular. Estos puntos son llamados singulares, y deben ser incluidos en la lista para valores máximos y/o mínimos. Este tema (parametrización no suave) esta por fuera de los objetivos de este texto, vease [8].

Ejemplo 3.7 *Nos gustaria determinar el valor mínimo que puede alcanzar la expresión*

$$y = \sum_{i=1}^{n} a_i x_i^2$$

con respecto a las variables x_i ($a_i > 0$ son numeros dados) bajo la restricción

$$c = \sum_{i=1}^{n} x_i,$$

donde c es una constante dada. Las condiciones de optimalidad nos llevan a

$$2a_j x_j + \lambda = 0, \quad j = 1, 2, \ldots, n,$$

de tal manera

$$x_j = -\frac{\lambda}{2a_j}$$

Si tomamos de nuevo estas expresiones en la restriccioñ

$$c = -\sum_{i=1}^{n} \frac{\lambda}{2a_j},$$

entonces

$$\lambda = -\frac{2c}{\sum_{i=1}^{n} \frac{1}{a_i}},$$

y en consecuencia

$$x_j = \frac{c}{a_j \sum_{i=1}^{n} \frac{1}{a_i}}$$

que es la unica solución al sistema. Note que dado que la función objetivo crece hacia $+\infty$ cuando algunas de las variables crecen indefinidamente, el que se alcance el mínimo esta garantizado. Por lo tanto la solución debe corresponder al mínimo. El mánimo es $+\infty$,dado que la restricción no es capaz de evitar que algunas de las variables crezcan de forma indefinida.

El siguiente ejemplo es mas sofisticado pero instructivo.

Ejemplo 3.8 *Sea a un vector fijo en $\mathbb{R}^3$. Se quiere determinar los valores extremos del funcional lineal ax bajo las restricciones.*

$$\begin{aligned} x_1x_2 + x_1x_3 + x_2x_3 &= 0, \\ |x|^2 = x_1^2 + x_2^2 + x_3^2 &= 1. \end{aligned}$$

Por simplicidad utilizaremos la notación

$$\det x = x_1x_2 + x_1x_3 + x_2x_3$$

Note que el conjunto de vectores x que satisfacen las dos restricciones es un subconjunto de la esfera unidad en $\mathbb{R}^3$, por lo tanto es acotado, y el funcional de costo necesariamente alcanza sus valores máximo y mínimo. Estos pueden ser detectados examinando las condiciones de optimalidad, explicitamente

$$\begin{aligned} a + \lambda_1 Ax + \lambda_2 x &= 0 \\ \det x &= 0, \\ |x|^2 &= 1, \end{aligned}$$

La matriz A es

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

La primera ecuación (vectorial) nos informa que el vector a debe ser una combinación lineal (cuyos coeficientes son los multiplicadores) de los vectores Ax y x. Eliminando los multiplicadores, podemos escribir esta ecuación de forma equivalente pidiendo

$$0 = \begin{vmatrix} a \\ Ax \\ x \end{vmatrix},$$

Pero dado que

$$0 = \begin{vmatrix} a \\ x \\ x \end{vmatrix}$$

tambien tenemos

$$0 = \begin{vmatrix} a \\ Ax \\ x \end{vmatrix} + \begin{vmatrix} a \\ x \\ x \end{vmatrix} = \begin{vmatrix} a \\ x + Ax \\ x \end{vmatrix}$$

Sea $e = (1, 1, 1 =$. Note que $x + Ax = \begin{pmatrix} x & e \end{pmatrix} e$, asi que

$$0 = xe \begin{vmatrix} a \\ e \\ x \end{vmatrix}.$$

Notese que xe no se puede anular porque para uno de nuestros vectores factibles x tenemos que

$$|xe|^2 = |x|^2 + 2\det x = 1.$$

Por lo tanto debemos tener

$$\begin{vmatrix} a \\ e \\ x \end{vmatrix} = 0,$$

y esto implica

$$x = sa + te$$

para ciertos coeficientes s y t. Si tomamos esta expresión en las restricciones $|x|^2 = 1$ *y* $\det x = 0$*, despues de algunos calculos obtenemos las dos ecuaciones cuadraticas*

$$\begin{aligned} \det(a)s^2 + 2stae + 3t^2 = 0, \\ |a|^2 s^2 + 2stae + 3t^2 = 1. \end{aligned}$$

inmediatamente obtenemos

$$s^2(|a|^2 - \det a) = 1.$$

Note que (¿por qué?)

$$|a|^2 - \det a \geq 0$$

Si

$$|a|^2 - \det a = 0,$$

entonces la ecuación para s es inconsistente. De hecho, en esta situción no tenemos ninguna solución para las condiciones de optimalidad. Aunque esta situación puede ocurrir solo cuando a es un multiplo de e (¿Por qué?), asi que el funcional de costo es

$$cxe,$$

como se indico anteriormente

$$|xe|^2 = |x|^2 + 2\det x = 1,$$

y por lo tanto el funcional es continuo para todos los vectores factibles.

Suponga que a es tal que

$$|a|^2 - \det a > 0,$$

En este caso tenemos dos soluciones para s:

$$s = \pm \frac{1}{\sqrt{|a|^2 - \det a}}$$

Tomando estos valores en la primera de las ecuaciones cuadraticas (3.2) y resolviendo para t (usando de nuevo la formula para $(ae)^2$*) nos lleva a*

$$t = \pm \frac{1}{3} \pm \frac{ae}{3\sqrt{|a|^2 - \det a}}$$

El par de soluciones validas son

$$\left(\frac{1}{\sqrt{|a|^2-\det a}}, \pm\frac{1}{3}-\frac{ae}{3\sqrt{|a|^2-\det a}}\right), \quad \left(\frac{-1}{\sqrt{|a|^2-\det a}}, \pm\frac{1}{3}+\frac{ae}{3\sqrt{|a|^2-\det a}}\right)$$

Entre estos cuatro vecores tenemos que en encontrar los valores máximo y mínimo del producto interno

$$ax = s|a|^2 + tae,$$

Examinado las cuatro posiblidades, concluimos que el valor máximo es

$$\frac{1}{3}\left(2\sqrt{|a|^2-\det a}+|ae|\right),$$

y se alcanza en

$$x = sa + te$$

para

$$s = \frac{1}{\sqrt{|a|^2-\det a}}, \quad t = \frac{ae}{3|ae|} - \frac{ae}{3\sqrt{|a|^2-\det a}}$$

El valor mínimo es el máximo cambiado en signo y se alcanza en el punto opuesto del máximo.

3.3. Condiciones de optimalidad Karush-Kuhn-Tucker

Nos gustaria tratar el caso general en el cual algunas de las restricciones estan en la forma de igualdades y otras estan en la forma de desigualdades:

$$\text{Minimizar} \quad f(x) \quad \text{sujeto a} \quad g(x) \leq 0, \quad h(x) = 0.$$

Exploremos primero que clase de condiciones necesita satisfacer un punto de tal manera que pueda ser una solución óptima a nuestro problema. ¿Qué tiene de especial tal punto?

Sea $x^{(0)}$ un tal punto mínimo, y sea M el conjunto de índices $M = \{1, 2, \ldots, m\}$ donde m es precisamente el numero de componentes de g. Consideramos el siguiente subconjunto de M:

$$J = \{j \in M : g_j(x^{(0)}) = 0\}.$$

Puede suceder que este conjunto es el conjunto vacío. Para j perteneciente a $M \backslash J$, decimos que la restricción correspondiente esta inactiva. Miremos al problema auxiliar

$$\text{Minimizar} \quad f(x) \quad \text{sujeto a} \quad g_j(x) \leq 0, \quad j \in J, \quad h(x) = 0.$$

Nuestra solución inicial $x^{(0)}$ sera ciertamente un punto mínimo local, aunque tal vez no global (¿Por qué?), y en consecencia dado que todas las restricciones

de este problema estan en la forma de igualdades existen una coleccioñ de multiplicadores

$$\mu_j, \quad j \in J, \quad \lambda \in \mathbb{R}^d,$$

tal que

$$\nabla f(x^{(0)}) + \sum_{j \in J} \mu_j \nabla g_j(x^{(0)}) + \lambda \nabla h(x^{(0)}) = 0. \tag{3.2}$$

Mas aun, afirmamos que μ_j puede ser tomado no negativo. La razon intuitiva de esto es que f tiene un mínimo en el punto $x^{(0)}$ pero cada una de las g_j tiene un máximo, ya que $g_j(x^{(0)}) = 0$ es ek valor máximo que puede tomar g_j en el conjunto factible para nuestro problema inicial. Por lo tanto los gradientes de f y g_j en el punto $x^{(0)}$ deben "Apuntar en direcciones diferentes". Esta afirmacion requiere naturalmente mas rigor y cuidado, pero es suficiente para nuestros propositos. Para $j \in M \backslash J$, tomamos $\mu_j = 0$. Asi llegamos a las condiciones necesarias de optimalidad, conocidad como las condiciones de Karush-Kuhn-Tucker (KKH).

Teorema 3.9 *Si x es una solución óptima no singular para nuestro problema, entonces existe un vector de multiplicadores (μ, λ) tal que*

$$\begin{aligned} \nabla f(x) + \mu \nabla g(x) + \lambda \nabla h(x) &= 0, \\ \mu g(x) &= 0, \\ \mu \geq 0, \quad g(x) \leq 0. \quad h(x) &= 0. \end{aligned}$$

La necesidad de tales condiciones quiere decir que las soluciones óptimas de deben buscar entre aquellos vectores x para los cuales podemos encontrar un conjunto de multiplicadores (μ, λ) que satisfacen estas condiciones. Esta información nos deje selecionar aquelos puntos que son factibles para ser puntos mínimos. En aquellas situaciones en las cuales todas aquellas soluciones pueden encontrarse, y tenemos la informacion que nuestron problema debe de hecho tener al menos alguna solución, estas puden ser identificados simplemente calculando el costo de todos los candidatos y decidiendo el mínimo.

Antes de analizar varios ejemplos explicitos, nos gustaria hacer un par de observaciones importantes.

1. Si el problema de optimización consisten en encontrar el máximo en vez del mínimo, jugando con los signos menos no es dificil escribir los cambios en las condiciones KKT. De hecho tendriamos

$$\begin{aligned} \nabla f(x) + \mu \nabla g(x) + \lambda \nabla h(x) &= 0, \\ \mu g(x) &= 0, \\ \mu \leq 0, \quad g(x) \leq 0. \quad h(x) &= 0. \end{aligned}$$

   Si no nos preocupammos acerca del signo de los componentes de μ, la lista de puntos factibles para extremos crecera significativamente con soluciones que no pueden ser ni máximos ni mínimos, dado que los signos de los componentes de μ estaran mesclados, positivos y negativos son no negativos,

Por lo tento, si todos los componentes de μ sonb no negativos el punto correspondiente puede ser un punto de mínimo (nunca un máximo), si todas las componentes son no positivas el punto puede sr un punto de mínimo (nunca de máximo), y si hay componentes positivas y negativas de μ entonces el punto nunca pude ser de máximo o mínimo (punto de silla).

2. Las condiciones
$$\mu \geq 0, \quad g(x) \leq 0, \quad \mu g(x) = 0,$$
son equivalentes (¿Por qué?) a
$$\mu \geq 0, \quad g(x) \leq 0, \quad \mu_j g_j(x) = 0, \quad j = 1, 2, \ldots, m.$$
asi que para encontrar todas las soluciones para las condiciones KKT tenemos que encontrar todas las soluciones al sistema de $n+m+d$ ecuaciones con $n + m + d$ incognitas x, μ, $lambda$,
$$\begin{aligned} \nabla f(x) + \mu \nabla g(x) + \lambda h(x) &= 0, \\ \mu_j g_j &= 0, \quad j = 1, 2, \ldots, m. \\ h_i &= 0, \quad i = 1, 2, \ldots, d \end{aligned}$$
que tambien satisfacen $\mu \geq 0$, $g(x) \leq 0$, en el caso que estemos interesados en el mínimo, y que satisfacen $\mu \leq 0$, $g(x) \leq 0$, en el caso de un máximo. Cuando se trata de resolver el sistema anterior, y gracias a su estructura particular, podemos simpre proceder examinando los 2^m casos
$$\mu_i = 0, \quad g_j(x) = 0, \quad j \in J, i \in M \setminus J,$$
donde J recorre todos los posibles subconjuntos de $M = \{1, 2, \ldots, m\}$. Entre todas las diferentes posibilidades que podemos obtener, debemos descartar aquellas que son imposibles o en las que no estamos interesados. Aunque puden haber otros metodos razonables para resolver el sistema, esta es la forma racional de organizar los calculos. Por otro lado, cualquier observacion basada en la naturaleza particular del problema, que nos lleve a descartar algunas de las soluciones, puede simplificar considerablemente el procedimiento que nos lleva a la solución.

Veamos varios ejemplos.

Ejemplo 3.10 ** Suponga que cierta red electrica consiste en tres diferentes canales a traves de los cuales para cierta corriente electrica. Si x_i, $i = 1, 2, 3$ es la cantidad de energía que pasa a traves del canal i, la perdida total en la red esta dada por la función*
$$p(x_1, x_2, x_3) = x_3 + \frac{1}{2}\left(x_1^2 + x_2^2 + \frac{x_3^2}{10}\right).$$
Si se planea transfirar una cantidad total r, determine las cantidades a traves de cada canal para minimizar la perdida de energía.

Evidentemente el problema se reduce a encontrar el mínimo de la función anterior proviendo una medida de la perdida de energía bajo las restricciónes

$$x_1 + x_2 + x_3 = r \quad x_i \geq 0.$$

Dado que el conjunto de puntos en $\mathbb{R}^3$ *que satisface estas restricciones es acotado (¿Por qué?) el punto mínimo que estamos buscando debe ser una de las soluciones del sistema de optimalidad*

$$x_1 - \mu_1 + \lambda = 0, \quad x_2 - \mu_2 + \lambda = 0, \quad 1 + \frac{x_3}{10} - \mu_3 + \lambda = 0,$$
$$\mu_1 x_1 = 0, \quad \mu_2 x_2 = 0, \quad c_1 + x_2 + x_3 = r,$$

donde las incognitas son $x_1, x_2, x_3, \mu_1, \mu_2, \mu_3, \lambda$. *Estamos interesados en aquellas soluciones que satisfacen*$x_1, x_2, x_3, \mu_1, \mu_2, \mu_3 \geq 0$. *Note que* λ *esta asociado con la restricción de igualdad* $x_1 + x_2 + x_3 = r$, *y como tal no podemos tener ninguna restricción en su signo. Resolver el sistema anterior requiere un poco de abilidad en encontrar todas las soluciones correspondientes a las ocho posibilidades.*

$$\begin{aligned}
x_1 = x_2 = x_3 = 0,\\
x_1 = x_2 = \mu_3 = 0,\\
x_1 = \mu_2 = x_3 = 0,\\
\mu_1 = x_2 = x_3 = 0,\\
x_1 = \mu_2 = \mu_3 = 0,\\
\mu_1 = x_2 = \mu_3 = 0,\\
\mu_1 = \mu_2 = x_3 = 0,\\
\mu_1 = \mu_2 = \mu_3 = 0,
\end{aligned}$$

Despues de estudiar con cierto cuidado estas posibilidad y descartar las soluciones en las cuales no estamos interesados, llegamos a la solución óptima

$$(r/2, r/2, 0) \quad \textit{con multiplicadores asociados} \quad (0, 0, 1 - r/2, -r/2)$$

cuando $r \leq 2$ *y*

$$((10+r)/12, (10+r)/12, 5(r-2)/6) \quad \textit{con multiplicadores} \quad (0, 0, 0, -(10+r)/12)$$

para $r \geq 2$. *Notese como ambas soluciones coinciden cuando* $r = 2$.

Ejemplo 3.11 *Tenemos una cuerda de longitud a para atar una caja de arriba a abajo a traves de dos direcciones perpendiculares. ¿Cúal es el volumen máximo que puede tener una tal caja?*

Nos gustaria determinar el máximo de la función volumen

$$V = (x_1, x_2, x_3) = x_1 x_2 x_3$$

sujeto a las condiciones

$$x_1, x_2, x_3 \geq 0, \quad 2x_1 + 2x_2 + 4x_3 \leq a,$$

asumiendo que x_3 es la altura. Con algun cuidado de los signos menos que deben ser introducidos para transformar nuestro problema al formato estándar, tenemos el sistema

$$x_2x_3 - \mu_1 + 2\mu 4 = 0, \quad x_1x_3 - \mu_2 + 2\mu_4 = 0, \quad x_1x_2 \quad \mu_3 + 4\mu_4 = 0,$$
$$x_1\mu_1 = 0, x_2\mu_2 = 0, \quad x_2\mu_3 = 0, \quad (2x_1 + 2x_2 + 4x_3 - a)\mu_4 = 0.$$

Buscamos aquellas soluciones con

$$\begin{aligned} x_1, x_2, x_3 &\geq 0, \\ 2x_1 + 2x_2 + 4x_3 &\leq a, \\ \mu_1, \mu_2, \mu_3, \mu_4 &\leq 0. \end{aligned}$$

Dado que las primeras tres ecuaciones nos permiten expresar μ_1, μ_2 y μ_3 en terminos de x_1 ,x_2, x_3 y μ_4, podemos eliminar las primeras variables y obtener un sistema equivalente

$$\begin{aligned} x_1(x_2x_3 - 2\mu_4) &= 0, \\ x_2(x_1x_3 - 2\mu_4) &= 0, \\ x_3(x_2x_1 - 4\mu_4) &= 0, \\ (2x_1 + 2x_2 + 4x_3 - a)\mu_4 &= 0. \end{aligned}$$

Si sumamos las tres primeras ecuaciones y tenemos en cuenta la ultima llegamos a

$$== 3x_1 + x_2 + x_3 + a\mu_4,$$

por lo tanto

$$\mu_4 = -3x_1x_2x_3/a.$$

Si aplicamos esta indentidad a las ecuaciones del sistema anterior y notamos que el mámixmo de V no puede anularse ($x_1x_2x_3 \neq 0$), desde que este es el mínimo, obtenemos la siguiente solución unica para el máximo

$$x_1 = x_2 = a/6, \quad x_3 = a/12.$$

Mas aun los multiplicadores asociados son

$$(0, 0, 0, -a^2/144),$$

Que efectivamente corresponden a un punto máximo, Dado que la region de la restricción es acotada esta es la solución óptima buscada.

Ejemplo 3.12 *Nos gustaria encontrar el mínimo y el máximo de*

$$f(x_1, x_2, x_3) = x_1^3 + x_2^3 + x_3^3$$

sobre la region determinada por las restricciones

$$x_1^2 + x_2^2 + x_3^2 \leq 4, \quad x_1 + x_2 + x_3 \leq 1.$$

Es un ejercicio sencillo escribir las condiciones KKT para esta situación.

$$\begin{aligned} 3x_1^2 + \mu_1 2x_1 + \mu_2 &= 0, \\ 3x_2^2 + \mu_1 2x_2 + \mu_2 &= 0, \\ 3x_3^2 + \mu_1 2x_3 + \mu_2 &= 0, \\ \mu_1(x_1^2 + x_2^2 + x_3^2 - 4) &= 0, \\ \mu_2(x_1 + x_2 + x_3 - 1) &= 0, \\ x_1^2 + x_2^2 + x_3^2 &\leq 4, \\ x_1 + x_2 + x_3 &\leq 1. \end{aligned}$$

Ademas, debemos tener en cuenta las restricciones en los signos de los multiplicadores, μ_1 y μ_2, cuando estamos buscando el máximo o el mínimo: $\mu_1, \mu_2 \geq 0$ si estamos buscando el máximo y $\mu_1, \mu_2 \leq 0$ si estamos buscando el mínimo. Separamos la discución el sistema anterior en cuatro casos.

1. *$\mu_1 = \mu_2 = 0$: en este caso inmediatamente obtenemos la solución $x_1 = x_2 = x_3 = 0$, que es admisible para ambos el máximo y el mínimo,*

2. *$\mu_1 = 0$, $x_1 + x_2 + x_3 = 1$: es directo obtener*

$$\mu_2 = -3x_1^2 = -3x_2^2 = -3x_3^2,$$

por lo tanto $(1/3, 1/3, 1/3)$ es la solución unica, que es admisible para el máximo pero no para el mínimo, dado que $\mu_2 = -1/3$;

3. *$\mu_2 = 0$, $x_1^2 + x_2^2 + x_3^2 = 4$: añadiendo las tres primeras ecuaciones (despues de eliminar μ_2) y teniendo en cuenta que la suma de los cuadrados es la unidad, obtenemos*

$$12 + 3(x_1 + x_2 + x_3)\mu_1 = 0.$$

Esta identidad descarta la posibilidad que $x_1 + x_x + x_3 = 0$, asi

$$\mu_1 = \frac{-6}{x_1 + x_2 + x_3}.$$

Si aplicamos esta igualdad a las tres primeras condiciondes de optimalidad, llegamos a la conclusión que o las corrdenadas de los puntos se anulan o de lo contrario su valor comun es

$$\frac{4}{x_1 + x_2 + x_3}$$

Ahora encontramos las siguientes soluciones (descartando simultaneamente aquellas que nosatisfacen $x_1 + x_2 + x_3 \leq 1$):

$$(-2,0,0), \quad (0,-2,0), \quad (0,0,-2),$$
$$(-\sqrt{2},-\sqrt{2},0) \quad (-\sqrt{2},0,-\sqrt{2}) \quad (0,-\sqrt{2},-\sqrt{2})$$
$$\left(\frac{-2}{\sqrt{3}},\frac{-2}{\sqrt{3}},\frac{-2}{\sqrt{3}}\right).$$

Dado que todas las soluciones satisfacen $x_1 + x_2 + x_3 < 0$ ($\mu_1 > 0$), estas seran factibles para el mínimo.

4. *El ultimo caso esta asociado con las igualdades*

$$x_1 + x_2 + x_3 = 1, \quad x_1^2 + x_2^2 + x_3^2 = 4$$

que ha sido resuelta en la sección anterior. Despues de calcular los valores de la funcion costo f en todos aquellos puntos seleccionados llegamos a la conclusión que el máximo se alcanza en

$$\left(\frac{1}{3}+\frac{\sqrt{22}}{6},\frac{1}{3}+\frac{\sqrt{22}}{6},\frac{1}{3}+\frac{\sqrt{22}}{3}\right),$$

y el valor mínimo se alcanza en

$$(-2,0,0), \quad (0,-2,0), \quad (0,0,-2).$$

Compare estos resultados con aquellos de los ejemplos 3.5 y 3.6

3.4. Convexidad

Hemos desarrollado un entendimiento basico de como detectar las condiciones necesarias de optimalidad que deben sr satisfechas en un punto de máximo o mínimo (local).

Tambien hemos enfatizado suficientemente en el hecho que las soluciones a las condiciones KKT pueden incluir otros puntos que no son necesariamente los puntos que estamos buscando. Pueden existir soluciones a las condiciones de optimalidad KKT que no corresponden a los valores extremos. La pregunta fundamental que nos gustaria hacer es si hay algun requisito en la función objetivo y/o las funciones que expresan las restricciones de tal manera que podamos asegurar que las soluciones de las condiciones KKT son exactamente los puntos donde se alcanza el mínimo (o el máximo), sin una discusión "a posteriori" sobre la naturaleza de las diferentes soluciones. Como veremos despues, este tema es muy importante dado que en la mayoria de situaciones que uno encuentra en la practica, las soluciones de las condiciones KKT no pueden ser encontradas explicitamente y necesitan ser aproximadas. Uno nunca esta seguro que todas las soluciones se han encontrado. Dada la relevancia de este tema, analizareos

la situacion en mayor detalle comenzando con el caso mas basico de programación no lineal de tal manera que podames entender la idea de la nocion de convexidad.

Consideremos el siguiente problema

$$\text{Minimizar} \quad f(x) \quad x \in \mathbb{R}$$

asumiendo que f es tan regular como lo necesitemos. La condicion KKT se reduce en esta situación simplificada a

$$f'(x) = 0.$$

Esta ecuación (no lineal) puede tener varias soluciones, incluso infinitas, y cualquiera de ellas podria ser el púnto mínimo buscado. Pero algunos de esos pueden corresponder a puntos de mínimo local, máximo local o puntos silla. Consideremos la siguiente situación:

$$f(x) \to +\infty \quad \text{when} \quad x \to \pm\infty$$
$$f'(x) = 0 \quad \text{tiene solución unica} \quad x_0.$$

No es dificil notar que x_0 es realmente el punto de mínimo global, y por lo tanto la solución unica a nuestro problema inicial (¿Por qué?). El primer requisito en los limites infinitos de f puede ser relativamente sencillo de revisar una vez conocemos f. ¿Pero como podemos estar seguros de la unicidad de la solución de la ecuación los puntos criticos sin necesidad de resolverla? Si suponemos que f admite una segunda derivada continua $f''(x)$, y $f''(x) > 0$ para toda x (que en muchos casos es facilmente verificable), entonces podemos asegurar el segundo requisito, o sea la ecación $f'(x) = 0$ solo puede tener una solución ('?Por qué?)

Ademas, sabemos del calculo elemental que la condición $f'' > 0$ quiere decir que f es convexo. En resumen.

1. $f(x) \to +\infty$, $x \to \pm\infty$: la ecuación $f'(x) = 0$ tiene al menos una solución que corresponde al máximo global de f.

2. $f''(x) > 0$ para todo x: f es estrictamente convexo, y la ecuación $f'(x) = 0$ tiene a lo mas una solución.

Por lo tanto, si se cumplen ambos requerimientos, la unica solución de $f'(x) = 0$ la unica solución de la ecuación $f'(x) = 0$ correspondera al mínimo global. Basicamente esta es la razon porque la convexidad es deseable cuando se trata con problemas de minimización (lo mismo ocurre con concavidad para los problemas de maximización). Otra forma de resumir las observaciones anteriores es que bajo convexidad , las condiciones necesarias para para la optimalidad se convierten en suficientes, dado que la convexidad elimina la existencia de mínimos locales que no son globales. Dado que estamos ante uno de los conceptos mas importantes en optimización, vamos a verlo con mas coudado.

Definición 3.13 *Un conjunto K en $\mathbb{R}^n$ es convexo si para cada par de vectores $x, y \in K$, el segmento de linea que los conecta tambien está en K:*

$$tx + (1-t)y \in K, \quad t \in [0,1]$$

Una función $f : K \subset \mathbb{R}^n \to \mathbb{R}^n$ es convexa si K es un conjunto convexo y

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y), \quad \textit{cuando} \quad x.y \in K, \quad t \in [0,1].$$

Antes de intetar un mejor entendimiento de esta condición, vale la pena convencerse que es un concepto clave relevante a problemas de minimización. Aprenderemos mas de convexidad en la siguiente sección. Nuestros lectores probablemente saben que quiere decir la condición de convexidad geometricamente. Vease la figura 3.2

La razon por la cual la convexidad es tan importante el los problemas de minimizacioñ puede ser formulada como sigue.

Teorema 3.14 *Sea*

$$f : K \subset \mathbb{R}^n \to \mathbb{R}^n$$

convexa, donde K tambien es convexa. Si x_0 es un mínimo local para f entonces x_0 es un mínimo global para f en K.

Para justificar este resultado, note que la condicion que x_0 sea un mínimo local para f en K quiere decir

$$f(x_0) \leq f(x), \quad |x - x_0| < \epsilon, x \in K,$$

donde $\epsilon > 0$ es un numero dado. Definimos

$$B_\epsilon = \{x \in K : |x - x_0| < \epsilon\}$$

Sea $y \in K$ un punto arbitrario (en K). El segmento que une a x_0 y y tiene puntos que pertenecen a B_ϵ,

$$tx_0 + (1-t)y \in B_\epsilon, \quad t \text{ suficientemente cerca de 1.}$$

Por lo tanto para tal t y por la convexidad de f,

$$f(x_0) \leq f(tx_0 + (1-t)y) \leq tf(x_0) + (1-t)f(y)$$

Reorganizando estos terminos tenemos

$$(1-t)f(x_0) \leq (1-t)f(y)$$

Dado que $(1-t) > 0$ para tal t concluimos que

$$f(x_0) \leq f(y).$$

La arbitrariedad de $y \in K$ nos da el resultado esperado.

La definición de convexidad (Insertar link aqui) quiere decir que para cada pareja de puntos en K, x, y , los valores de f en el segmento que los une no sobrepasa los de la "'línea"'que pasa por $(x, f(x))$, $(y, f(y))$. Dicho de manera diferente los valores de f estan debajo de cada una de sus secantes. Si la función f es diferenciable, entonces se puede dar una caracterización alternativa de la convexidad. Esto es mas apropiado en nuestro contexto porque se puede relacionar directamente a las condiciones de optimalidad. Incluso, si la función f es dos veces diferenciable con matriz hessiana continua, entonces la convexidad se puede dar en terminos de las segundas derivadas.

Proposición 3.15 *Sea*

$$f : K \subset \mathbb{R}^n \to \mathbb{R}$$

una función continua donde K es convexo abierto y convexo.

1. *Si f es diferenciable y ∇f es continua, entonces f es convexa si y solo si*

$$f(y) \geq f(x) + \nabla f(x)(y - x), \quad x, y \in K. \tag{3.3}$$

2. *Si f es dos veces diferenciable y $\nabla^2 f$ es continua, entonces f es convexa si y solo si $\nabla^2 f(x)$ es semidefinida positiva para todo $x \in K$.*

Dado que la expresión

$$f(x) + \nabla f(x)(y - x),$$

considerada como una función de $y \in K$, es la ecuación del hiperplano tangente a f en x, la desigualdad (3.3) dice que la grafica de f esta por encima de cualquiera de sus hiperplanos tangentes. Para la caracterización con las segundas derivadas, podemos decir de manera equivalente que si f es dos veces diferenciable y $\nabla^2 f$ es continuaentonces f es convexa si y solo si los valores propios de $\nabla^2 f(x)$ son no negativos en cada $x \in K$.

La prueba de la proposición 3.15 procede en dos pasos. Primero, mostraremos que la afirmación es cierta para funciones de una variable $h : J \to \mathbb{R}$ donde J es un intervalo en $mathbbR$.

Por lo tanto supongamos que h es diferenciable y que satisface

$$h(tx + (1 - t)y) \leq th(x) + (1 - t)h(y), \quad x, y \in J, \quad t \in [0, 1].$$

Reorganizando y manipulando estos terminos podemos transformar la ecuación anterior (cuando $1 - t > 0$ y $x \neq y$) en

$$(y - x)\frac{h(tx + (1 - t)y) - h(x)}{(1 - t)(y - x)} \leq h(y) - h(x)$$

Dado que esta desigualdad es corecta para cada $t \in [0, 1]$, tomando limites cuando $t \to 1^-$, podemos concluir que

$$(y - x)h'(x) \leq h(y) - h(x).$$

Esta es la primera parte de la proposición. Si ademas asumimos que h tiene una segunda derivada en cada punto de J donde las desigualdades anteriores se cumplen, tendremos dependiendo si $y > x$ o $x > y$,

$$\frac{h(y) - h(x)}{y - x} \geq h'(x), \quad \frac{h(y) - h(x)}{y - x} \leq h'(x).$$

Por el teorema del valor medio aplicado a h' aseguramos la existencia de un z tal que

$$, \quad \text{o} \quad h'(z) \leq h'(x), z \leq x.$$

La arbitrariedad en y lleva a la arbitrariedad de z, y por lo tanto h' es una función no decreciente que se traduce en la no negatividad de h''. Este es el criterio de convexidad cuando las funciones son dos veces diferenciables.

Finalmente argumentamos que si h es una función dos veces diferenciable, y su segunda derivada es no negativa entonces necesariamente debemos tener

$$h(tx + (1-t)y) \leq th(x) + (1-t)h(y), \quad x, y \in J, \quad t \in [0,1].$$

Para lograr esto reorganizamos los terminos en la expresión

$$th(x) + (1-t)h(y) - h(tx + (1-t)y)$$

de la siguiente manera:

$$\begin{aligned} t(h(x) - h(tx + (1-t)y)) + (1-t)(h(y) - h(tx + (1-t)y)) \\ = t(1-t)(x-y)h'(a) - t(1-t)(x-y)h'(b), \end{aligned}$$

donde hemos usado el teorema del valor medio, y los puntos a y b estan entre x y $tx + (1-t)y$, y $tx + (1-t)y$ y y, respectivamente. De nuevo por el teorema del valor medio aplicado a h' existe un número c entre x y y tal que

$$th(x) + (1-t)h(y) - h(tx - (1-t)y) = t(1-t)(x-y)(a-b)h''(c).$$

Si notamos que el producto (x-y)(a-b) es siempre no negativo (aunque ambos factores pueden ser negativos), y ya que los otros factores son negativos nos lleva a tener

$$th(x) + (1-t)h(y) - h(tx + (1-t)y) \geq 0$$

como se deseaba.

Para el segundo paso, considere una función $f : K \to \mathbb{R}$ de varias variables. La prueba para este caso esta basada en lo que ya hemos mostrado para funciones de una variable simplemente aplicando las conclusiones previas a las secciones

$$h(s) = f(x + s(y - x))$$

para x, y fijos, y aplicando convenietemente la regla de la cadena para calcular las derivadas y segundas derivadas. Dejamos los detalles al lector interesado.

A veces es importante, porque nos lleva a consecuencias significativas, saber lo que ocurre cuando cambiamos las desigualdes por desigualdades estrictas en las tres formas de comprobar convexidad. Las funciones que tienen esta propiedad adicional se llaman estrictamente convexas, hablaremos ahora de convexidad extricta.

Definición 3.16 *Una función $f : K \to \mathbb{R}$, donde $K \subset \mathbb{R}^n$ es convexo se llama estrictamente convexa si f es continua (esto es redundante) y*

$$f(tx - (1-t)y) < tf(x) + (1-t)f(y), \quad x \neq y \in K, \quad t \in)(0,1)$$

Se puede mostrar una caracterización simililar a la de la propisición 3.15 para convexidad estricta cuando la función f es mas regular.

Proposición 3.17 *Sea*

$$f : K \subset \mathbb{R}^n \to \mathbb{R}$$

una función continua donde K es convexo y abierto.

1. *Si f es diferenciable y ∇f es continua, entonces f es estrictamente convexa si y solo si*

$$f(y) > f(x) + \nabla f(x)(y - x), \quad x \neq y \in K;$$

2. *Si f es dos veces diferenciable y $\nabla^2 f$ es continua, entonces f es estrictamente convexo si y solo si su matriz Hessiana es definida positiva en cada punto en K, o incluso, por el criterio de Sylvester, los subdeterminantes principales son estrictamente positivos para cada punto de K.*

La prueba es un ejercicio interesante de revisar la pruba de la propisición 3.15 revisando las desigualdades y las desigualdades estrictas. Como una regla general podemos decir que la convexidad sin convexidad estrica esta asociada tipicamente con "las partes planas de la grafica"

Terminamos esta sección revisando varios ejemplos de funciones convexas.

Ejemplo 3.18 *Toda funcion linear (o afín) es convexa pero no estrictamente convexa.*

Hay cuatro operaciones elementales que respetan convexidad:

1. *Una combinación linal de funciones convexas con coeficientes no negativos es una función convexa.*

2. *Si $T : \mathbb{R}^n \to \mathbb{R}$ es lineal y $h : \mathbb{R} \to \mathbb{R}$ es convexa, entonces la composición $f(x) = h(Tx)$ es convexa.*

3. *Si $g : K \subset \mathbb{R}^n \to \mathbb{R}$ es convexa y $h : \mathbb{R} \to \mathbb{R}$ es convexa y no decreciente, entonces la composición $f(x) = h(Tx)$ es convexa.*

4. *El supremo de cualquier familia de funciones convexas es de nuevo una función convexa.*

Es facil verificar estas cuatro afirmaciones usando directamente la definicion de convexidad.

Usando funciones lineales y las operaciones basicas mencionadas anteriormente podemos generar nuevas funciones convexas o revisar la convexidad de ejemplos conocidos. Por ejemplo si notamos que

$$|x| = \sup\{ax : |a| = 1\},$$

resulta que la distancia al origen, $|x|$ es una función convexa. Mas aun dado que

$$h(t))t^p, \quad t \geq 0,$$

es una función convexa si $p \geq 1$,

$$g(x) = |x|^p$$

es convexa si $p \geq 1$. Esta funcion es estrictamente convexa si $p > 1$. Si tomamos

$$h(t) = \sqrt{1+t^2},$$

que es de nuevo una funcion convexa no decreciente cuando $t \geq 0$, la función

$$g(x) = \sqrt{1+|x|^p}$$

sera convexa si $p \geq 2$, y estrictamente convexa si $p > 2$. Las funciones

$$|x|^p + |x|^q, \quad ax + \sqrt{1+|x|^p},$$
$$|x-a|^p, \quad |ax|^p + |bx|^q,$$

son convexas cuando los exponentes p, q son mayores o iguales a 1.

Si f es una función que no es convexa, definimos su convexificación como

$$Cf(x) = \sup\{h(x) : h \leq f, h \quad convexa\}$$

que es la función convexa mas grande entre aquellas que estan debajo de f. Si f ya es convexa, su convexificacion es f misma. Poe ejemplo, si f esta dada por

$$f(x) = \min\{(x+1)^2, (x-1)^2\},$$

que no es convexa, su compactificación es la función definida a trozos por

$$g(x) = \begin{cases} (x+1)^2 & x \leq -1, \\ 0 & |x| \leq 1 \\ (x-1)^2 & x \geq 1 \end{cases}$$

una función convexa pero no estrictamente convexa.

!!!

Ejemplo 3.19 *Considerese ahora*

$$Minimizar \quad |x|^2 bajo \quad ax \leq c, \quad bx \leq d,$$

con $a, n, x \in \mathbb{R}^n$, $c, d \in \mathbb{R}$. De nuevo la función objetivo es estrictamente convexa y aquellas en las restricciones son lineales, asi que el problema tiene a lo mas una solución, la cual en caso que exista, debe ser la unica solución para las condiciones KKT. Estas son

$$2x + \mu_1 a + \mu_2 b = 0,$$
$$\mu_1(ax - c) = 0, \quad \mu_2(bx - d) = 0,$$

junto con $\mu_1, \mu_2 \geq 0$, $ax \leq c$, $bx \leq d$. Un estudio completo de estas ecuaciones llevara a cuatro posibilidades

1. $\mu_1 = \mu_2 = 0$, $x = 0$: *Esta sera la solución óptima si* $x = 0$ *es factible i.e,* $c \geq 0$, $d \geq 0$.

2. $\mu_1 = 0$, $\mu_2 = -2d/|b|^2$, $x = (d/|b|^2)b$: d *debe ser negativo, y la factibilidad de este vector* x *implica la restricción*

$$dab \leq c|b|^2;$$

3. $\mu_2 = 0$, $\mu_1 = -2c/|a|^2$, $x = (x/|a|^2)a$: c *debe ser negativo, y la factibilidad de* x *implica*

$$cab \leq d|a|^2;$$

4. *Cuando ambos multiplicadores no se anulan, pueden ser determinados como la solución del sistema lineal*

$$\begin{aligned} |a|^2\mu_1 + ab\mu_2 &= -2c, \\ ab\mu_1 + |b|^2\mu_2 &= -2d, \end{aligned}$$

Cuyo determinante, $|a|^2|b|^2 - (ab)^2$ *no se anula a menos que* a *y* b *sean colineales. La solución esta dada por*

$$\mu_1 = \frac{2(dab - c|b|^2)}{|a|^2|b|^2 - (ab)^2}, \quad \mu_2 = \frac{2(cab - d|a|^2)}{|a|^2|b|^2 - (ab)^2}$$

y el vector óptimo es

$$x = \frac{1}{2}(\mu_1 a + \mu_2 b).$$

Para resumir, dependiendo de los datos particulares a, b, c, d, *podemos tener las siguientes cuatro situaciones:*

1. $c \geq 0$, $d \geq 0$;

2. $d < 0$, $dab \leq c|b|^2$;

3. $c < 0$, $cab \leq d|a|^2$;

4. $dab > c|b|^2$, $cab > d|a|^2$.

Es importante anotar que estas no son cuatro soluciones diferentes sino una unica que depende de la relación entre los diferentes vectores a *y* b, *y los escalares* c *y* d. *Todas estas posibilidades estas dibujadas en la figura 3.3.*

Ejemplo 3.20 *Considere el problema de encontrar el mínimo de*

$$|x|^4 + |x - a|^2$$

bajo la restricción

$$|x|^2 \leq 1,$$

donde a es un vector dado. Dado a la convexidad estricta de la función costo, la solución óptima debe corresponder a la solución unica de las condiciones KKT

$$4|x|^2x + 2(x-a) + \mu 2x = 0, \quad \mu(|x|^2 - 1) = 0,$$

donde debemos tener en cuanta las restricciones adicionales $\mu \geq 0$, $|x| \leq 1$. *Los dos casos admisibles son*

$$\mu = 0, \quad x = t_a, \quad 2|a|^2t^3 + t - 1 = 0,$$

$$\mu = |a| - 3, \quad x = a/|a|.$$

En la primera situación debemos exigir $|a| \leq 3$ *para que* $x = ta$ *sea factible, dado que*

$$(2|x|^2 + a)x = a.$$

Note que el polinomio cubico que especifica a t tiene una raiz real unica que esta en el intervalo $(0, 1/|a|)$ *si* $|a| \leq 3$, *si* $|a| > 3$,*la solución óptima corresponde a la segunda alternativa.*

Ejemplo 3.21 *Se planea disear una estructura apuntalada como la mostrada en la figura 3.4 de acuerdo al criterio del peso mínimo sujeto a la restricción de la máxima deflección permitida en el nodo libre y la coma mínima en las secciones transversales de los miembros. Los datos de este problema son*

$$a_1, a_2, A_0, A_1, A_2, x_0.$$

Todos deben ser positivos y dependen de la geomtria de la estructura, constantes del material, cargas en los puntos indicados, etc. Especificamente el problema puede ser puesto como

$$\textit{Minimizar} \quad a_1x_1 + a_2x_2$$

sujeto a

$$\frac{A_1}{x_1} + \frac{A_2}{x_2} \leq A_0, \quad x_1, x_2 \geq 0,$$

donde x_1 *y* x_2 *son precisamente las areas transversales que se planean disear.*

Se invita al lector a verificar que este PPNL es convexo, asi que la solución óptima se puede entontrar resolviendo las condiciones KKT. Especificamente, su $\mu_i, i = 1, 2, 3$, *son los multiplicadores asiciados con las tres restricciones en la forma de desigualdades, tenemos*

$$
\begin{aligned}
a_1 - \frac{\mu_1 A_1}{x_1^2} - \mu_2 &= 0,\\
a_2 - \frac{\mu_2 A_2}{x_2^2} - \mu_3 &= 0,\\
\mu_1 \left(\frac{A_1}{x_1} + \frac{A_2}{x_2} - A_0 \right) &= 0,\\
\mu_2(x_0 - x_1) &= 0,\\
\mu_3(x_0 - x_3) &= 0,\\
\frac{A_1}{x_1} + \frac{A_2}{x_2} - A_0 &\leq 0,\\
x_0 - x_1 \leq 0, \quad x_0 - x_2 &\leq 0,\\
\mu_1 \geq 0, \quad \mu_2 \geq 0, \quad \mu_2 &\geq 0.
\end{aligned}
$$

Una discución completa de la solución requerira un numero diferente de casos dependiendo de los valores particulares del conjunto de datos anteriores. Para definir el problema tomaremos

$$
\begin{aligned}
a_1 = \frac{1}{5}, \quad a_2 &= \tfrac{1}{6}\\
A_0 = 12, \quad A_1 = 25 \quad A_2 &= 100,\\
x_0 = 10 &
\end{aligned}
$$

Todos estos dados en sus unidades correspondientes. Para este conjunto de datos, la solución óptima es

$$x_1 = 10, \quad x_2 = \sqrt{120}$$

Con multiplocadore

$$\mu_1 = \mu_2 = \mu_3 = 0, \quad \mu_2 = \frac{1}{5}.$$

Los detalles se le dejan al lector interesado.

3.5. Dualidad y convexidad

Como en PL, le podemos asociar a cada PPNL otro PPNL, llamado su dual, de tal manera que hay una relación cercana entre los dos. Dado que la PNL es mucho mas complicada que su contraparte líneal, podemos esperar que la dualidad en PNL es mucho mas complicada. Esta sección pretende ser una mera introducción al tema. Desde el punto de vista practico, la dualidad en PNL aparece como una herramienta poderosa en tratar de aproximar de mejor manera las soluciones óptimas en PNL. Como tal, esta cercanamente conectada a la convexidad como veremos.

Definición 3.22 *Dado un problema primario*

$$\textit{Minimizar} \quad f(x) \quad \textit{bajo} \quad g(x) \leq 0, \quad h(x) = 0,$$

definimos su cual como el PPNL

$$\textit{Maximizar}\theta(\mu, \lambda) \quad \textit{bajo}\mu \geq 0,$$

donde a θ se le llama la función dual, y esta definida en los pares de multiplicadores (μ, λ) de la siguiente manera

$$\theta(\mu, \lambda) = \inf_x [f(x) + \mu g(x) + \lambda h(x)].$$

El porque el problema esta definido de esta manera de volvera mas claro a medida que entendamos mejor la conección entre estos dos PPNL y los relacionemos con las condiciones de optimalidad KKT. En cierto sentido, la idea de fondo es incorporar la condiciones necesarias para la optimalidad como parte de la factibilidad para un nuevo problema como sigue:

$$\text{Minimizar} \quad F(x, \mu, \lambda) = f(x)$$

sujeto a

$$g(x) \leq 0, \quad h(x) = 0, \nabla f(x) + \mu \nabla g(x) + \lambda \nabla h(x) = 0, \mu \geq 0, \quad \mu g(x) = 0.$$

La función que aparece en la definición de la función dual es conocido como el Lagrangiano asociado con el problema

$$L(x, \mu, \lambda) = f(x) + \mu g(x) + \lambda h(x).$$

Lema 3.23 *Suponga que las funciones f, g y h sol tales que el infimo que define a la función dual θ siempre se alcanza para todos los pares (μ, θ), $\mu \geq 0$. Sea $X = X(\mu, \theta)$ el punto donde se alcanza el mínimo de tal manera que*

$$\theta(\mu, \lambda) = f(X) + \mu g(X) + \lambda h(X).$$

Entonces si la función $X(\mu, \lambda)$ es diferenciabla, tambien lo es θ y

$$\nabla_\mu \theta(\mu, \lambda) = g(X), \nabla_\lambda(\mu, \lambda) = h(X).$$

Nuestra justificación es un calculo directo. Si

$$\theta(\mu, \lambda) = f(X) + \mu g(X) + \lambda h(X).$$

por un lado por la regla de la cadena,

$$\nabla_\mu \theta = (\nabla f(X) + \mu \nabla g(X) + \lambda \nabla h(X)) \nabla_\mu X + g(X);$$

pero por otro, si el Lagrangiano alcaza su mínimo en X, su gradiente con respecto a x debe anularse,

$$\nabla f(X) + \mu \nabla g(X) + \lambda \nabla h(X) = 0,$$

asi que

$$\nabla_\mu \theta = g(X)$$

como se deseaba. Tenemos un resultado simiplar con el gradiente respecto a λ

De la misma manera como argumentamos en el caso de PL, la dualidad se muestra en dos pasos. La siguiente proposición es normalmente conocida como dualidad debil.

Proposición 3.24 *Sean f, g y h diferenciables.*

1. *Siempre se tiene que*

$$\text{máx}\{\theta(\mu, \lambda) : \mu \geq 0\} \leq \text{mín}\{f(x) : g(x) \leq 0, h(x) = 0\}.$$

2. *Si (μ, λ) es factible para el problema dual ($\mu \geq 0$), x es factible para el primal ($g(x) \leq 0, h(x) = 0$), y*

$$\theta(\mu, \lambda) = f(x),$$

entonces (μ, λ) y x son soluciones óptimas para ambos el dual y el primal respectivamente.

La explicación es elemental. Note que su $\mu \geq 0$, $g(x) \leq 0$, y $h(x) = 0$, entonces

$$\theta(\mu, \lambda) \leq f(x) + \mu g(x) + \lambda h(x) \leq f(x).$$

Esto implica dualidad debil. La segunda parte de la farmación tambien es directa.

La diferencia

$$\text{mín}\{f(x) : g(x) \leq 0, h(x) = 0\} - \text{máx}\{\theta(\mu, \lambda) : \mu \geq 0\}.$$

se le llama la difencia de dualidad. Cuando no hay tal diferencia, ambos problemas son equivalentes, y el primal puede ser resuelto resolviendo el dual. Esta es la idea principal de todos los algoritmos númericos para calcular las soluciones óptimas revisando el dual. Aparte de la interpretación del problema dual por si mismo, esta es la razon principar porque el problema dual es tan importante en PNL. La convexidad es de nuevo la hipotesis principal bajo la cual la diferencia de dualidad se anula.

Teorema 3.25 *Suponga que f y g son funciones convexas diferenciables, h es afín, y el problema de optimización que define la función dual es siempre soluble. Entonces ambos problemas, el primal y el dual, son solubles simultaneamente y no hay diferencia de dualidad.*

Para justificar, identifiquemos por (P) y (D) el primal y el dual respectivamente. Asuma primero que el primal es soluble, asi que existe un vector x y multiplicadores (μ, λ) que satisfacen las condiciones KKT, explicitamente,

$$\mu g(x) = 0, \quad \mu \geq 0, \quad f(x) \leq 0, \quad h(x) = 0.$$

,

Todas estas condiciones implican que x es factible para (P), y (μ, λ) es factible para (D); por la convexidad de f y g y por la linealidad de h, x es un punto donde se alcanza el mínimo del Lagrangiano, pero dado que $\mu g(x) = h(x) = 0$, tenemos

$$\theta(\mu, \lambda) = f(x).$$

La proposición 3.24 implica que (μ, λ) es una solución óptima para (D).

Conversamente, suponga que (μ, λ) es uuna solución óptima para el dual. Si entonces aplicamos las condiciones KKT a este problema obtenemos

$$\begin{aligned} \nabla_\mu \theta(\mu, \lambda) - y = 0, \quad \nabla_\lambda \theta(\mu, \lambda) = 0, \\ \mu \geq 0, \quad y \geq 0, \quad y\mu = 0, \end{aligned}$$

donde y es el multiplicador asociado con la restricción $\mu \geq 0$. Teniendo en cuenta el lema 3.23, si x es un punto donde

$$\theta(\mu, \lambda) = f(x) = \mu g(x) + \lambda h(x).$$

de tal manera que

$$\nabla f(x) + \mu \nabla g(x) + \lambda \nabla h(x) = 0,$$

y podemos reinterpretar esas condiciones de optimalidad como

$$\begin{aligned} g(x) = \nabla_\mu \theta(\mu, \lambda) = y \geq 0, \\ h(x) = \nabla_\lambda \theta(\mu, \lambda) = 0, \\ \mu g(x) = \mu y = 0. \end{aligned}$$

Bajo las suposiciones de convexidad para f, g y la linealidad en h que satisfacen ls condiciones KKT asegura que x es una solución óptima para (P). Es relevante enfatizar lo que esta dualidad significa que el mínimo

$$\text{mín}\{\text{máx}\{f(x) + \mu g(x) + \lambda h(x) : \mu \geq 0\} : g(x) \leq 0, h(x) = 0\}$$

y el máximo

$$\text{máx}\{\text{mín}\{f(x) + \mu g(x) + \lambda h(x) : g(x) \leq 0, h(x) = 0\} : \mu \geq 0\}$$

son iguales. La dualidad es siempre una pregunta de cuando las operaciones min-max son reversibles. Terminamos esta sección calculando la función dual en un ejemplo particular.

Ejemplo 3.26 *Considere el PPNL*

$$\textit{Minimizar} \quad x_1^2 + x_2^2 + x_3^2$$

sujeto a

$$x_1^2 + x_2^2 + 3x_3 \leq \frac{5}{2}, \quad x_1 + x_2 + x_3 = -2$$

Dado que en esta situción se dan los criterios de optimalidad, encomrar la función dual para este problema se reduce a resolver las condiciones KKT para (μ, λ) *fijos olvidando las constantes, i.e. resolver el sistema*

$$\begin{aligned} 2x_1 + 2x_1\mu + \lambda &= 0, \\ 2x_2 + 2x_2\mu + \lambda &= 0, \\ 2x_3 + 3\mu + \lambda &= 0, \end{aligned}$$

La solución una es

$$x_1 = x_2 = \frac{-\lambda}{2(1+\mu)}, \quad x_3 = \frac{-1}{2}(\lambda + 3\mu).$$

Llevando estos valores al lagrangiano correspondiente tenemos la función dual

$$\theta(\mu, \lambda) = \frac{-\lambda^3}{2(1+\mu)} - \frac{1}{4}(\lambda + 3\mu)^2 + \frac{5}{2}\mu + 2\lambda.$$

Ahora es cuestion de calcular que las soluciones óptimas para el dual y el primal estas relacionadas a traves de la dualidad y las condiciones KKT. Estas soluciones óptimas son

$$x_1 = x_2 = -\frac{1}{2}, \quad x_3 = -1, \quad \mu = \frac{1}{4}, \quad \lambda = \frac{5}{4},$$

y el valor óptimo para el primal y el dual es valor común es $3/2$.

3.6. Ejercicios

1. Determine los puntos criticos de la función

   $$f(x_1, x_2) = (2 - x_1 - x_2)^2 + (1 + x_1 + x_2 - x_1x_2)^2.$$

   E intente dicernir su naturaleza.

2. Encuentre el mínimo y el máximo de la función

   $$\sigma^2 = \sum_{i=1}^{n} T_i^2 x_i^2$$

   con respecto a las variables x_i sujeto a

   $$c = \sum_{i=1}^{n} x_i.$$

   c y T_i son constantes fijas.

3. Dado la función objetivo

$$P = (x_i - 1)^2 + x_n^2 + \sum_{i=1}^{n-1}(x_{i+1} - x_k)^2,$$

encuentre los puntos criticos de P

$a)$ sin ninguna restricción

$b)$ bajo

$$c = \sum_{i=1}^{n} a^i x_i,$$

donde c y a son constantes.

4. Muestre que la función

$$f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1 - x_2 - x_3$$

es convexa. Encuentre los valores extremos de f bajo las condiciones

$$x_1^2 + x_2^2 = 4, \quad -1 \leq x_3 \leq 1.$$

5. Describa la region del plana determinada por las desigualdades

$$x_1^2 - x_2^2 \leq 1, \quad x_1^2 + x_2^2 \leq 4.$$

Encuentre los valores extremos de

$$f(x_1, x_2) = x_1^2 + 2x_2^2 + x_1 x_2$$

sobre aquella region.

6. Defina las funciones

$$F(a) = \text{mín}\{f(x_1, x_2, x_3) : g(x_1, x_2, x_3) = a\},$$
$$G(a) = \text{mín}\{f(x_1, x_2, x_3) : g(x_1, x_2, x_3) \leq a\},$$
$$,$$

para $a \in \mathbb{R}$ y

$$f(x_1, x_2, x_3) = x_1^4 + x_1^2(1 - 2x_2^2) + x_2^2 + x_3^2 - 2x_1 + 1,$$
$$g(x_1, x_2, x_3) = x_1^4 + x_2^4 + x_3^4$$

Encuentre expresiones explicitas para $F(a)$ y $G(a)$. Estudie los dos problemas

$$\text{mín}\, F(a), \quad \text{mín}\, G(a)$$

y su relación. ¿Que puede concluir acerca del mínimo de la función f sobre todo $\mathbb{R}^3$? Haga lo mismo para escogencias mas simples de la función g.

7. Encontre el valor máximo y mínimo de

$$f(x_1, x_2) = \int_{x_1}^{x_2} \frac{1}{1+t^4} dt$$

sobre la región determinada por $x_1^2 x_2^2 = 1$.

8. Encuentre el punto mas cercano de la superficie $xy+xz+yz = 1$ al origen. Haga lo mismo para la superficie con ecuación $x^2 + y^2 - z^2 = 1$.

9. Resuelva el problema de función de utilidad Cobb-Douglas en el capítulo 1.

10. Resuelva el problena de la localización de varis puntos de servicio donde los clientes son conocidos (capítulo 1) para el siguiente conjunto de datos: Las localizaciones de los consumidores son

$$(1,0), (2,1), (-1,2), (3,-1), (-1,-2), (3,-2),$$

y se van a construir tres puntos de servicio.

11. Resuelva el problema de la escalera del capítulo 1.

12. Resuelva el problema del andamiaje del capítulo 1.

13. Cierto conjunto de datos experimentales que relaciona dos variables x y y esta a nuestra disposición.

$$(x_i, y_i), \quad i = 1, \ldots, n.$$

Se espera una relación lineal entre x y y, pero esto tipicamente no es posible de manera exacta. Determine los mejores coeficientes a, b de tal manera que el error cuadratico del conjunto de datos con respecto al modelo lineal

$$y = ax + b$$

sea mínimo.

14. Cuando cierto sistema $Ax = b$ no es soluble, podemos estar interesados en el vector x "mas cerca" de ser una solución minimizando el error cuadratico (o de otro tipo)

$$\text{Minimizar} \quad \frac{1}{2}|Ax - b|^2$$

sobre todos los posibles x. Resuelva este problema en general y aplique su solución al caso particular

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix}.$$

15. Se diseara una barra conica truncada, sujeta en su borde superior y colgando verticalmente (Vease ejercició 10 capítulo 1) bajo las siguientes restricciones (Figura 3.5): longitud total L, volumen total a nuestra disposición V, densidad del material que se va a usar ρ, Modulo de Young del material que se va a usar E. El problema es minimizar la elongación total bajo la acción de un peso dado W en su extremo inferior y su propio peso. Asumimos que la ley de Hooke es valida: Si la sección transversal a una distancia x del extremo superior se mueve a $y(x)$ bajo la acción de W, entonces el esfuerzo en tal sección es $y'(x)$ es proporcional (Con constante $1/E$) al "stress" en tal secció, y este "stress" es ademas es el cociente de la carga total actuando en el y el área correspondiente de la sección transversal. Se determinaran los radios R y r de las secciones superiores e inferiores.

    Si las secciones tranversales son cuadrados, de tal manera que la barra es una piramide trucada, ¿Se espera la misma solución?

Capítulo 4

Tecnicas de Aproximación

4.1. Introducción

Es probable que nuestros lectores hallan notado que resolver un problema de optimización explicitamente no es una tarea sencilla. De hecho, en la mayoria de casos es imposible. No solo para aquellos problemas con un número alto de variables para los cuales es imposibles de calcular a mano sus soluciones óptimas, sino tambien para muchos problemas de tamaño modesto para los cuales es imposible resolver y manipular tantas ecuaciones. Por lo tanto es importante mostrar la manera como se pueden aproximar solciones para problemas de optimización de manera eficiente. Esta necesidad es inevitable desde el punto de vista practico y de ingenieria, dado que las aproximaciones explicitas y precisas de kas soluciones son tan importantes como el entendimiento del problema de fondo. Como es usual, nuestra meta en este capítulo es cubrir el algoritmos basicos que los investigadores han usado a traves de los años para aproximar soluciones para problemas de PPNL sin tratar de agotar todas las posibilidades, describir los avances mas recienteso incluso mostar de donde vienen los algoritmos y porque tienen aquella estructura particular. Trataremos de motivar sin embargo los mas populares de manera que el lector tenga la intuición de su naturaleza sin entrar en los detalles tecnicos. Tambien es cierto que esta es un área muy tecnica que evoluciona rapidamente, asi que los métodos que parecen ser mejores ahora probablemente seran abondados en algunos años y reemplazados por viejas ideas con un nuevo enfoque o por tecnicas totalmente innovadoras. Vease por ejemplo [17] para una buena investigación en este tema y la importancia de los s de punto interior hoy en dia.

Hay tambien otra razon importante de porque no describimos una lista completa de algoritmos o entramos en detalles mas profundos, y esta es que actualmente existen paquetes de software muy poderosos cuya función es aproximar soluciones óptimas en una variedad de situaciones. Estas herramientas liberan al usuario de la necesidad de preocuparse demasiado acerca de los tecnicismos relacionadas con los algoritmos practicos y enfocarse en los asuntos de mode-

lamiento del problema y la interpretación e implementación de las soluciones. Algunos de estos paquetes de software son AIMMS, AMPL, AMSL, GAMS, Matlab (Optimization Toolbox) y SNOPT. Hay mucha información sobre este tema esparcida en Internet. Recomendamos especialmente el sitio [29], que es casi un sitio maestro para optimización. Particularmente cuando se esta buscanque que software es mejor para nuestras necesidades particulares, este es el sitio a visitar.

Para aquellos lectores interesados en profundizar su entendimiento de las tecnicas de aproximación y que esten dispuestos a seguir en esta dirección deberian definitivamente recurrir a alguna de las referencias que hemos seleccionado al final del texto. En particular, algoritmos para problemas grandes de optimización requieren considerable experiencia, paciencia y estudio. Aproximación es mucho mas que el contenido de este capítulo. Para fuentes completas vease [1], [13], [16], [22], [30] y [32].

Muchos métodos de aproximación númerica para PPNL tienen una naturaleza iterativa. Esto quiere decir que el proceso de aproximación procede en aproximaciones sucesivas cada vez mejores a las soluciones buscadas. De esta manera, tal algoritmo debe especificar un mecanismo para construir una nueva iteración de una (o varias) de las que ya tenemos. De esta manera construimos una sucesión $\{x_k\}$ de aproximaciones sucesivas a la solución óptima real x, confiando en que x_{k+1} estara mas cerca a x qye su aproximación precesente x_k. Es mas, un exponente $\alpha > 0$ caracteriza los algoritmos numericos eficiente cuando

$$|x_{k+1} - x| \leq C|x_k - x|^\alpha, \quad C > 0.$$

Si interpretamos la diferencia

$$||x_j - x||$$

como una medida del error que cometemnos cuando tomamos x_j como x, la desigualdad anterior dice que en cada nueva iteración el error es reducido en una potencia con exponente α. Mientras mas grande es α mejor el método, dado que el error disminuye mas rapidamente, pero el calculo de la nueva aproximación puede ser mas caro en terminos computacionales. Debe haber siempre un balance entre la eficiencia de un algoritmo y el costo computacional asociado con su implementación.

Cada algoritmo número especifico debe determinar de manera precisa la forma de pasar de una iteración x_k a la siguiente x_{k+1}. Dado que estos son vectores, podemos pensar esto proceso como una estrategia de dos pasos:

1. Dirección de busqueda: Escoger el vector d_k que apunta hacia x_{k+1} desde x_k, de tal manera que d_k es paralelo a $x_{k+1} - x_k$; a este vector d_k se le llama la dirección de busqueda.

2. Tamaño del paso: Una vez se ha escogido d_k se trata de determinar un parametro t_k de tal manera que

$$x_{k+1} = x_k + t_k d_k$$

a t_k se le llama el tamaño del paso.

Estos dos elementos, d_k y y_k son suficientes para determinar una nueva iteración desde una anterior. Cada algoritmo de optimización completo debe indicar e incorporar estos dos ingredientes: dirección de la busqueda y tamaño del paso.

Restringiremos nuestra atención en primera instancia a los PPNL sin restricciónes, y trataresmos en este contexto algunos de los algoritmos para decidir en el tamanó del paso y en la dirección de la busqueda. En el primer caso, describiremos brevemente los algoritmos de tamanó del paso fijo, interpolación y regla de oro, ademas nos enfocaremos en los métodos de decenso mas rápido, gradiente conjugado y aquellos cuasi Newton para el segundo caso. Mas adelante nos concentraremos en las ideas basicas para trabajar con PPNL con restricciónes incluyendo penalizaciónes y barreras, el método dual y finalmente el método del Lagrangiano aumentado. Trataremos de permanecer tan especificos en cada situación como sea posible, con la idea de no confundir al lector con demasiados resultados.

4.2. Métodos de busqueda lineal

Comenzaremos estudiando la aproximación númerica del problema

$$\text{Minimizar} \quad g(x) \quad x \in \mathbb{R}^n$$

Como se menciono en la sección anterior, la busqueda del tamanó del paso debe hacerse una vez se ha determinado una dirección de busqueda. Por lo tanto supondremos que se ha escogidouna dirección de busqueda d_k que parte de la aproximación x_k, y nos gustaria escoger un parametro t_k de tal modo que

$$x_{k+1} = x_k + t_k d_k$$

sera nuestra proxima iteración a la solución óptima que buscamos. Si estamos interesados encontrar el mínimo global de $g(x)$, donde asumimos que g esta definida en todo $\mathbb{R}^n$, y que el correspondiente problema de minimización tiene solución, nuestra mejor apuesta es escoger t_k como la solución del problema

$$\text{Minimizar} \quad g(x_k + td_k), \quad t \in \mathbb{R}.$$

Si definimos

$$h(t) = g(x_k + td_k), \quad t \in \mathbb{R}.$$

esto nos lleva a considerar el problema de minimización unidimensional de la función h. Dado que h es una sección unidimensional de g a este paso se le conoce normalmente como una busqueda lineal. Por lo tanto nos concentramos en encontrar una aproximación para

$$\text{Minimizar} \quad h(t), \quad t \in \mathbb{R}$$

Procederemos a describir brevemente algunos de los métodos mas simples que pueden y son usados para la situación uno dimensional. Ciertamente podriamos

tratar de determinar de manera exacta el punto mínimo para la función h, pero esta estrategia no es tentadora desde un punto de vista practico, dado que si pudieramos encontrar de manera exacta el punto mínimo, el hecho que tendremos que resolver estos problemas muchas veces de manera sistematica para problemas de dimiensiones superiores para encontrar una manera eficiente de aproximar el valor mínimo para problemas unidimensionales.

Tamaño del paso fijo o variable *Este es el mas elemental de todos los métodos. Consiste en escoger un tamanó del paso fijo t y tomando aproximaciones sucesivas poniendo*

$$t_{i+1} - t_i = t.$$

Continuamos de esta manera hasta que

$$h(t_{i+1}) \geq h(t_i);$$

en este punto reducimos el tamaño de t en un factor significativo y cambiamos su signo. Ahora procedemos tomando

$$t_{i+1} - t_i = t$$

con este nuevo valor de t hasta que de nuevo obtenemos

$$h(t_{i+1}) \geq h(t_i).$$

Continuamos repitiendo este proceso iterativamente hasta que una presición preasignada se obtenga.

Interpolación *El método de interpolación consiste en interpolar los valores de h en tres puntos t_1, t_2, t_3, por un polinimio cuadratico cuyo mínimo es facilmente calculable. Es conveniente para que la ubicación relativa de los valores cumpla.*

$$\begin{aligned} t_1 < t_2 < t_3, \\ h(t_2) < h(t_1), \\ h(t_2) < h(t_3). \end{aligned}$$

Tal punto mínimo, $t*$ que minimiza al polinomio de interpolación se considera una buena aproximación al punto mínimo real de h. Si se desea mayor presición, $t*$ remplaza a alguno de los t_i de tal manera que la posición relativa de los valores de h cumple las condiciones anteriores, y se hacen de nuevo los calculos hasta que se alcance cierto umbral de presición. Este método resulta muy eficiente para funciones suaves, especialmente porque existen formulas esplicitas para $t*$ en terminos de t_i y de $h(t_i)$. De hecho,

$$t* = \frac{1}{2} \frac{t_1^2(h(t_2) - h(t_3)) + t_2^2(h(t_3) - h(t_1)) + t_3^2(h(t_1) - h(t_2))}{t_1(h(t_2) - h(t_3)) + t_2(h(t_3) - h(t_1)) + t_3(h(t_1) - h(t_2))}$$

–+Notese que este método es exacto en una sola iteracion cuando h es cuadratica.

Método de la sección dorada *El algoritmo de la sección dorada trata de reducir el tamanó del intervalo donde se encuentra el mínimoa* $k = 0{,}618034..$ *Este factor es la sección dorada, y es la solución positiva de la ecuación* $1 + k = 1/k$. *Supongamos que el mínimo se alcanza en un punto en el intervalo. Consideramos los puntos*

$$t_3 = kt_1 + (1-k)t_2, \quad t_4 = kt_2 + (1-k)t_1.$$

Comparando los valores de $h(t_3)$ *y* $h(t_4)$ *procedemos como sigue:*

1. *si* $h(t_3) > h(t_4)$, *el mínimo esta en el intervalo* (t_3, t_2), *de manera que* t_3 *reemplaza a* t_1.

2. *si* $h(t_3) < h(t_4)$, *entonces el mínimo pertenece al intervalo* (t_1, t_4) *y* t_4 *toma el lugar de* t_2.

Se aplica este procedimiento iterativamente hasta que el intervalo tiene una longitud menor que un valor umbral preasignado.

Método de Fibonnaci *Esta es una variación del método anterior en la cual la razon k depende de la iteración que estamos calculando. Los valores de k cambian de acuerdo con la suseción de Fibbonaci, que se define recursivemente como*

$$a_1 = a_2 = 1, \quad a_{j+2} = a_{j+1} + a_j, \quad j \geq 1.$$

En cada iteración usamos el parametro

$$k_j = \frac{a_j}{a_{j+1}}$$

en vez de k. Note que k_j *tiende a k, la sección dorada cuando* $j \to \infty$
No es justo decir que un método particular es siempre mejor que cualquier otro. Dependiendo del problema, un método puede ser preferible sobre cualquiera de los otros. Cada usuario puede encontrar sus preferencias experimentando. En general podemos decir que el método de interpolación funciona bastante bien cuando la función es suave. Puede usarse la tecnica de la sección dorada, aunque para un algoritmo para el caso general existen metodos mas sofisticados. Veanse las referencias citadas en la introducción a este capítulo.

4.3. Metodos de gradiente

Una vez hemos tratado el tema de la escogencia del tamanó del paso cuando se haya escogido una dirección de busqueda, estudiaremos la escogencia de esta dirección. Esta es una pregunta mas compleja y fundamental. El exito de un método particular se ve altamente por un buen algoritmo para escoger estas direcciones de busqueda. Hay escencialmete dos grupos de algoritmos para escoger la dirección de busqueda. Aquelos que no requieren información de la

función objetivo, y aquellos basados en la información que obtienen del gradiente de la función objetivo. Dado que es razonable pensar que podemos usar la información que viene del gradiente para nuestro beneficio. Nos enfocaremos en los metodos de gradiente ya que estos son ampliamente usados, Describiremos tres tipos de metodos que estan entre las escogencias mas populares: el decenso mas empinado, el gradiente conjugado y los quasi-Newton.

La idea basica de todos estos algoritmos se basa en la noción de dirección de decenso para una función suave.

Definición 4.1 *(Dirección de decenso) Un vector $d \in \mathbb{R}^n$ es una dirección de decenso para una función suave $g : \mathbb{R}^n \to \mathbb{R}$ en un punto $x \in \mathbb{R}^n$ si*

$$\nabla f(x)d < 0$$

La razon de esta definición es bastante simple. Solo hay que notar que si definimos

$$h(t) = g(x + td),$$

entonces por la regla de la cadena

$$h'(0) = \nabla f(x)d.$$

Por lo tanto si d es una dirección de decenso esta derivada es negativa, y por lo tando los valores de g decrecen a medida que nos movamos por d desde x al menos localmente.

Cualquier metodo de gradiente esta asociado con diferentes formas de escoger la dirección de decenso. Como anunciamos anteriormente los metodos de gradiente conjugado y cuasi-Newton estan entre los mas populares y eficientes. Describiremos brevemente los metodos de decenso mas empinado quasi-Newton, pero profundizaremos un poco mas en los metodos de gradiente conjugado.

Decenso mas empinado *Es bien sabido que el gradiente de una función en un punto, $\nabla g(x)$, da la dirección en la cual la función incrementa de manera mas rapida. Por lo tanto $-\nabla g(x)$ nos da la dirección de "decenso mas rápido" desde el punto x. Entonces el gradiente de g es evaluado en cada iteración, y tomamos como direción de busqueda*

$$d_k = -\nabla g(x_k)$$

Notese que d_k es siempre una dirección de decenso, de hecho la dirección de decenso mas empinado dado que

$$\nabla g(x_k)d_k = -||\nabla g(x_k)||^2 \leq 0.$$

Si el gradiente se anula entonces x_k es el punto mínimo buscado. De lo contrario d_k es una dirección de decenso. Para resumir, el algoritmo del decenso mas empinado puede ser descrito como sigue:

1. ***Inicialización*** *Escoja x_0 la aproximación inicial.*

2. ***Dirección de busqueda*** *Dada una aproximación x_k defina $d_k = -\nabla g(x_k)$.*

3. ***Criterio de Parada*** *Para un valor umbral predeterminado $\epsilon > 0$ si*

$$||d_k|| \leq \epsilon,$$

detengase. La aproximación actual es lo suficientemente buena, i.e esta lo suficientemente cerca del verdadero punto mínimo. De lo contrario continúe. Notese que en tal punto mínimo el gradiente se anula.

4. ***Busqueda lineal*** *Encuentre una aproximación al problema de minimización*

$$\textit{Minimizar} \quad g(x_k + td_k), \quad t \in \mathbb{R}.$$

(vease la sección anterior). Sea t_k una aproximación a tal punto mínimo.

5. ***Nueva aproximación*** *Definase $x_{k+1} = x_k + t_k d_k$. Regrese al paso 2.*

Metodos Quasi-Newton *El metodo de Newton toma como dirección de busqueda*

$$d = -[\nabla^2 g(x)]^{-1} \nabla g(x),$$

donde

$$\nabla^2 g(x)$$

es la matrix (simetrica) Hessiana de g en x.

Proposición 4.2 *Si $\nabla^2 g(x)$ es definida positva, entonces d escogido de la forma anterior es una dirección de decenso para g en x. Es mas, si A es cualquier matriz definida positiva, la dirección*

$$-A\nabla g(x)$$

es una dirección de decenso de g en x.

La prueba es elemental y se le deja al lector. Mas importante que esto es entender de donde viene la dirección de decenso para el metodo de Newton. De la expansión de Taylor tenemos

$$g(x) \approx g(x_0) + \nabla g(x_0) + \frac{1}{2}(x - x_0)^T \nabla^2 g(x_0)(x - x_0),$$

Asi que por diferenciación

$$\nabla g(x \approx \nabla(x_0) + \nabla^2 g(x_0)(x - x_0).$$

Si x es un punto mínimo, debemos obligar que $\nabla g(x) = 0$, y esto lleva a

$$x = x_0 - [\nabla^2 g(x_0)]^{-1} \nabla g(x_0).$$

Esta es la explicación de la forma de la dirección de busqueda para el metodo de Newton

El calculo de la inversa de la matriz Hessiana es poco llamativo desde un punto de vista practico. Tratar de superar este problema lo lleva a uno a considerar los algoritmos cuasi-Newton donde la inversa de la matriz Hessiana es aproximada sucesivamente usando solamente el gradiente de g (las primeras derivadas). Simplemente daremos los dos algoritmos cuasi-Newton mas importantes sin mas justificación.

1. ***Inicialización.*** $x_1 = \mathbf{1}$, $p_1 = -H_1 \nabla g(x_1)$. *Aquí* $\mathbf{1}$ *es la matriz identidad.*
2. ***Busqueda lineal.*** *Encuentre una aproximación al problema de mínimización unidimensional*

$$Minimizar \quad g(x_k + tp_k), \quad t \in \mathbb{R}.$$

   Sea t_k *una aproximación a tal punto mínimo*

3. ***Nueva aproximación*** *Ponga* $x_{k+1} = x_k + t_k d_k$
4. ***Criterio de Parada*** *Para un valor umbral predeterminado* $\epsilon > 0$, *si*

$$||\nabla g(x_{k+1})|| \leq \epsilon,$$

   detengase: La aproximación actual x_{k+1} *es suficientemente buena, i.e suficientemente cerca del punto mínimo verdadero. En caso contrario continué.*

5. ***Nueva dirección de busqueda.*** *Sea*

$$q_k = \nabla g(x_{k+1}) - \nabla g(x_k), \quad p_{k+1} = -H_{k+1} \nabla g(x_{k+1}),$$

   y regrese al paso 2.

La forma de la matriz H_{k+1} *distingue algoritmos diferentes.*

1. ***Algoritmo de Davidon-Fletcher-Powell.*** *Tome*

$$H_{k+1} = H_k + t_k \frac{p_k p_k^T}{p_k q_k} - \frac{H_k q_k q_k^T H_k}{q_k^T H_k q_k}.$$

2. ***Algoritmo de Broyden-Goldfarb-Shanno*** *Tome*

$$H_{k+1} = t_k \frac{p_k p_k^T}{q_k p_k} + \left(\mathbf{1} - \frac{p_k q_k^T}{q_k p_k} \right) H_k \left(\mathbf{1} - \frac{q_k p_k^T}{q_k p_k} \right)$$

4.4. Metodos de Gradiente Conjugado

El algoritmo de gradiente conjugado es un procedimiento mas elaborado para decidir la dirección de busqueda comparado con el metodo del decenso mas empinado. Su interes puede ser motivado como una forma de solucionar el problema que la dirección dada por el metodo del decenso mas empinado no

es generalmente la mejor escogencia, y por lo tanto el metodo del decenso mas empinado puede ser extremadamente lento en acercarse al punto mínimo. El concepto principal asociado con el algoritmo de gradiente conjugado es aquel de direcciones conjugadas de una función cuadratica. En lo que sigue restringiremos nuestra atención, para motivar el origen del metodo por si mismo, a la función cuadratica.

$$g(x) = \frac{1}{2}x^T Ax - bx + c$$

donde A es una matriz definida positiva de tamaño $n \times n$ (de tal manera que el problema de minimización correspondiente tiene una solución unica), b es un vector, y c es una constante.

Definición 4.3 *Un conjunto de vectores p_i, $i = 1, 21 \ldots, n$, es un conjunto de direcciónes conjugadas para g si*

$$p_i^T A p_j = 0. \quad i \neq j, \quad p_i^T A p_i = \gamma_i, \quad i = 1, 2, \ldots, n.$$

En particular, este conjunto de vectores es linealmente independiente, y ellos forman una base para $\mathbb{R}^n$. Especificamos el interes de las direcciones conjugadas con respecto a los problemas de minimización en dos resultados. El primero establece la relevancia de las direcciónes conjugadas con respecto a los problemas de minimización. El segundo indica una de las varias posibilidades para determinar conjuntos de direcciones conjugadas dada una función cuadratica. Finalmente escribiremos el algoritmo de gradiente conjugado para una función objetivo general g, no necesariamente cuadratica.

Lema 4.4 *Sea p_j un conjunbto de direcciones conjugadas con respecto a la función cuadratica g mencionada anteriormente, y sea*

1. *x_1 arbitrario, $g_1 = -\nabla g(x_1)$;*
2. *para $k \geq 1$,*

$$t_k = -\frac{g_k p_k}{p_k^T A p_k}, \quad x_{k+1} = x_k + t_k p_k, \quad g_{k+1} = \nabla g(x_{k+1}) = g_k + t_k A p_k.$$

Para todo j tenemos En particular

$$g_{n+1} p_k = 0, \quad k = 1, 2, \ldots, n.$$

Esto implica que $g_{n+1} = 0$, y que por lo tanto g obtiene su mínimo en x_{n+1}

Este resultado nos dice que usando un conjunto de direccione conjugadas como direcciones de busqueda para un funcional cuadratico, podemos encontrar el mínimo en exactamente n pasos, donde n es la dimensión del problema. Es interesante ver de donde viene escogencia de t_k. Si consideramos la función

$$h(t) = g(x_k + t p_k),$$

resulta que el valor de t en el cúal se toma el mínimo de h es precisamente el que hemos escogido para t_k. Esto se deduce facilmente teniendo en cuenta que g es cuadratica y usando apropiadamente la regla de la cadena.

Para probar el Lema 4.4, notese que el mínimo que estamos buscando es la solución del sistema lineal

$$Ax - b = 0$$

dado que el gradiente de g es

$$\nabla g(x) = Ax - b$$

Por lo tanto pretendemos resolver el sistema lineal anterior en n pasos donde n es la dimensión de la matriz. Argumentando por inducción en el indice j. Asi suponemos que

$$g_{j+1}p_k = 0, \quad k = 1, 2, \ldots, j$$

y queremos concluir que

$$g_{j+2}p_k = 0. \quad k = 1, 2, \ldots, j+1$$

Por un lado tenemos

$$g_{j+2}p_k = (g_{j+1} + t_{j+1}Ap_{j+1} = p_k = 0$$

si $k = 1, 2, \ldots, j$ por hipotesis de inducción y el hecho que estamos trabajando con direcciones conjugadas. Por otro lado la identidad

$$g_{j+2}p_{j+1} = (g_{j+1} + t_j + 1Ap_{j+1})p_{j+1} = 0$$

sigue de la escogencia de t_{j+1}. Notese que la escogencia de t_k es dada por como determinamos el mínimo de la función cuadratica en t, $g(x_k + tp_k)$, como mencionamos anteriormente.

Una vez hemos visto el interes en los conjuntos de direcciones conjugadas, damos y justificamos una de las varias formas de construir recursivamente dichos conjuntos de direcciones y las aproximaciones sucesivas al punto mínimo simultaneamente.

Lema 4.5 *(Fletcher-Reeves) Sea*

1. $p_1 = -g_1 = -\nabla g(x_1)$;
2. *para* $k \geq 1$.

$$p_{k+1} = -g_{k+1} + \beta_k p_k, \quad \beta_k = \frac{|g_{k+1}|^2}{|g_k|^2},$$

donde g_j *es el gradiente de* g *en el punto* x_j. *El conjunto* $\{p_j\}$ *es un conjunto de direcciones conjugadas para* g

Para probar este resultado argumentamos de nuevo por inducción. Supongamos que hemos escogido k direcciones conjugadas p_j, $j = 1, 2, \ldots, k$, asi que segun el lema 4.4

$$g_{k+1}p_j = 0. \quad t_j^T Ap_j = g_{j+1} - g_j, \quad j = 1, 2, \ldots, k, \tag{4.1}$$

y

$$t_j = -\frac{g_j p_j}{p_j^T ap_j} = -\frac{g_j(-g_j + \beta_{j-1}p_{j-1})}{p_j^T Ap_j} = \frac{|g_j|^2}{p_j^T Ap_j},$$

dado que $g_j p_{j-1} = 0$. Si tomamos una neva dirección p_{k+1} como en el enunciado del lema, nos gustaria concluir que es conjugada con respecto a las anteriores . Teniendo en cuenta (4.1), tenemos.

$$\begin{aligned} p_{k+1}^T Ap_j &= (-g_{k+1} + \beta_k p_k)^T Ap_j \\ &= -g_{k+1}\frac{1}{t_j}(g_{j+1} - g_j) + \beta_k p_k^T Ap_j \\ = -\frac{g_{k+1}}{t_j}(p_{j+1} + \beta_j p_j + p_j - \beta_{j-1}p_{j-1}) &+ \beta_k p_k^T Ap_j. \end{aligned}$$

Para $j < k$ y asumiendo que t_j no se anula (de lo contrario $g_j = 0$. lo que implica que x_j es el punto mínimo, y que no hay necesidad de mas direcciones conugadas), vemos que esta expresión es cero por la hipótesis de inducción y el hecho que las direcciones escogidas anteriormente son conjugadas entre ellas. Para $j = k$, algunos de los terminos se anulan y otros no . Específicamente, teniendo en cuenta la formula para β_k obtenemos

$$\begin{aligned} p_{k+1}^T Ap_k &= \frac{g_{k+1}p_{k+1}}{t_k} + \beta_k p_k^T ap_k \\ &= \frac{p_k^T Ap_k}{(} - |g_{k+1}|^2 + \beta_k g_{k+1}p_k) + \beta_k p_k^t Ap_k \\ &= \frac{p_k^T Ap_k}{|g_k|^2}\left(-|g_{k+1}|^2 + \frac{|g_{k+1}|^2}{|g_k|^2}g_{k+1}p_k + |g_{k+1}|^2\right). \end{aligned}$$

Esta expresión se anula porque $g_{k+1}p_k = 0$.

Para una función objetivo general $g(x)$ el metodo de gradiente conjugado procede con aproximaciones sucesivas, y una nueva dirección de busqueda es escogida en cada iteración. La forma de estas direcciones de busqueda está justifificada por la discución anterior. De manera algoritmica podemos resumir el metodo del gradiente conjugado de la siguiente manera:

1. **Inicialización** Escoja x_0 la aproximación inicial.

2. **Criterio de Parada** Para un valor umbral predeterminado $\epsilon > 0$, si

$$||\nabla g(x_k)|| < \epsilon||,$$

detengase. La aproximación x_k es lo suficientemente buena, i.e, suficientemente cerca al verdadero punto mínimo. De lo contrario continue.

3. **Dirección de busqueda** Dada una aproximación x_k, defina $p_k = -\nabla g(x_k) + \beta_k p_{k-1}$, tomando p_{-1} de manera arbitraria.

4. **Busqueda lineal** Encuentre una aproximación al problema de minimizació sobre la linea (Vease sección 4.2)

$$\text{Minimizar} \quad g(x_k + tp_k), \quad t \in \mathbb{R}.$$

   Sea t_k una aproximación a tal punto mínimo.

5. **Nueva aproximación** Defina $x_{k+1} = x_k + t_k d_k$ y regrese al paso 2.

Las distintas variantes del algoritmo de gradiente conjugado corresponden a distintas maneras de tomar el parametro β_k. Las escogencias mas populares son las siguientes.

1. Algoritmo Feltcher-Reeves:

$$\beta_k = \begin{cases} 0, & k = jn, j = 0, 1, \ldots \\ \frac{||\nabla g(x_k)||^2}{||\nabla g(x_{k-1}||^2}, & \text{de lo contrario} \end{cases}$$

2. Algoritmo Polak-**Riviere**:

$$\beta_k = \begin{cases} 0, & k = jn, j = 0, 1, \ldots \\ \frac{\nabla g(x_k)(\nabla g(x_k) - \nabla g(x_{k-1}))}{||\nabla g(x_{k-1}||^2}, & \text{de lo contrario} \end{cases}$$

La razon para tomar $\beta_k = 0$ cada n iteraciones es para evitar el efecto de acumulación de errores numericos. Empiricamente, el algoritmo de Polak-Ribiere parece ser mas robusto.

4.5. Aproximación bajo restricciones

Dado que las restricciones son frecuentemente una parte escencial de los problemas de optimización, los algoritmos númericos para resolver o aproximar soluciones deben ser diseñados de tal manera que estos respeten las restricciones apropiadas y que lleven a la solución del problema de optimización sujeto a estas restricciones. Por lo tanto estamos interesados en describir los algoritmos que aproximan de forma eficiente las soluciones óptimas del PPNL.

$$\text{Minimizar} \quad f(x)$$

sujeto a

$$g(x) \leq 0, \quad h(x) = 0.$$

Hay escencialmente dos estrategias principales para tratar este problema numéricamente. O decidimos no usar multiplicadores, y por lo tanto no usar la información proveniente de las condiciones de optimalidad, o de lo contrario, tratamos

de utilizar esta información de alguna manera. La primera clase incluye las tecnicas de penalización y barreras. en la segunda categoria explicaremos el metodo dual estandar y el algoritmo Lagrangiano aumentado. En cualquier caso, los algoritmos estan construidos de tal manera que depende de alguna manera u otra del caso sin restricciones, asi que la idea detras de fondo es construir un problema relacionado pero sin restricciones y aplicarle a este alguno de los algoritmos que tenemos para problemas sin restricciones.

Penalización y Barreras *La idea detras de los metodos de penalización y barreras consiste en transformar el problema de optimización con restricciones de tal manara que los vectores no factibles estan prohibidos o al menos penalizados. Idealmente esto se logra considerando el siguiente problema de optimización.*

$$\textit{Minimizar} \quad f(x) + \hat{f}(x)$$

donde

$$\hat{f}(x) = \begin{cases} 0, & g(x) \leq 0, \quad h(x) = 0, \\ +\infty, & \textit{de lo contrario} \end{cases}$$

El efecto de añadir $\hat{f}(x)$ a f, como se puede ver en la definicion anterior es nulo si el vector es factible, asi que su costo es $f(x)$, como debe ser. Pero si x no es factible su costo es infinito, asi que es eliminado del proceso de minimización. Esto es exactamente lo que significan las restricciones. El nuevo problema no tiene restricciones. Sin embargo el hecho que el costo de $f + \hat{f}$ no es continua, dado que puede tomar el valor $+\infty$ de manera abrupta hace a este problema inapropiado para los algoritmos explicados en la sección anterior. Antes de aplicar tales algoritmos a este problema debemos hacer algo acerca de este costo especial $f + \hat{f}$.

Una posibilidad es no asignar un costo infinito a un vector no factible, sino simplemente penalizar de alguna manera a estos. Esta es la idea principal del metodo de penalización que es usada en muchas otras areas de las matemáticas. Si debemos satisfacer algunas restricciones que son dificiles de manejar las podemos ignorar. Pero aadimos una penalizacion al funcional de costo cuando estas no se cumplen para asi evitarlas. Una de las familias de penalizaciones más populares es

$$\hat{f}_r(x) = r \left(\sum_i \text{máx}\{0, g_i(x)\}^p + \sum_j |h_j(x)|^q \right),$$

donde $p, q > 1$ son exponentes y r es un parámetro de penalización. Note que si $p > 1$ la función $\text{máx}\{0, g_i(x)\}^p$ es diferenciable si g_i lo es. Mientras mas grande sea r es mas efectiva la penalización. Un caso típico que es usado a menudo es la penalización cuadratica que corresponde al caso $p = q = 2$. Aunque si bien es cierto que una penalización no prohíbe puntos no factibles , la proximación que se puede obtener usando el algoritmo de optimización no restringido en $f + \hat{f}$ puede no dar buenos resultados. Solamente cuando forzamos el parámetro r para que se vuelva cada vez mas grande, las aproximaciones correspondientes al los

problemas no restringidos se acercan cada vez mas a la solución del problema restringido.

Otra idea que puede ser usada cuando se trabaja con problemas restringidos es imitar la barrera infinita que impone $\hat{f}$ entre los conjuntos factibles y aquellos no factibles, pero de manera que esta no sea instantanea. Por ejemplo si tomamos

$$\hat{f}_r(x) = -\frac{1}{r}\sum_i \frac{1}{g_i(x)},$$

cuando las restricciones del problema vienen solamente de la forma de desigualdades del tipo $g(x) \leq 0$, vemos que a medida que nos movemos hacia la frontera del conjunto factible de tal manera que $g_i(x)$ se acerca cada vez mas a cero desde la parte negativa, la función $\hat{f}_r(x)$ crece cada vez mas hasta que eventualmente se vuelve infinita, poniendo una barrera al conjunto de los vectores factibles. El efecto del parametro r cuando este se vuelve grande es interferir tan poco como sea posible con el valor de la función objetivo f en el conjunto de los vectores factibles. Cuando $g_i(x) < 0$ y r es grande, entonces $1/r$ es pequeño de manera que $\hat{f}_r(x)$ tambien lo es. De nuevo solo cuando r es grande podemos obtener buenas aproximaciones aplicando el algoritmo no restringido al costo $f + \hat{f}_r$. A medida que $r \to +\infty$ estas aproximaciones tienden a la solución verdadera del problema original restringido.

Otra posibilidad interesante es tomar una barrera logaritmica del tipo

$$\hat{f}_r(x) = -\frac{1}{r}\sum_i \log(-g_i(x)).$$

Otra buena escogencia puede ser

$$\hat{f}_r(x) = \frac{1}{r}\sum_i \frac{1}{1 - r g_i(x)}.$$

En caso que algunad de las restricciones vengan en la forma de igualdades $h(x) = 0$ se puede añadir el termino

$$r^3 \sum_j \frac{h_j(x)^2}{1 - r^2 h_j(x)^2}$$

a $\hat{f}_r$,

Aunque las ideas anteriores son tentadoras por su simplicidad, en la practica debido al hecho que su efectividad esta ligada a valores altos del parametro r, aparecen errores númericos porque tenemos que manejar números muy grandes, y solo se obtiene una eficiencia moderada. Es mas, los resultados óptimos caundo se usan penalización o barreras se encuentran en los valores intermedios del parametro r, y este rango es altamente dependiente de cada problema particular, asi que se requiere una gran cantidad de experimentación para llegar a resultados óptimos. Por esta razon se han desarrollado algoritmos que tienen en cuenta los multiplicadores y optimalidad y/o las condiciones de dualidad. Restringiremos

nuestra atención al metodo dual estándar y terminaremos con una pequenña discución del metodo Lagrangiano aumentado.

Metodo dual *En el capítulo 3 aprendimos la relación entre un PPNL y su dual. De hecho si tenemos que resolver el problema primario*

$$\textit{Minimizar} \quad f(x)$$

suejeto a

$$g(x) \leq 0, \quad h(x) = 0,$$

sabemos que bajo hipótesis adecuadas podemos de manera equivalente tratar su dual

$$\textit{Maximizar} \quad \theta(\mu, \lambda)$$

bajo $\mu \geq 0$*, donde la función dual esta definida por*

$$\theta(\mu, \lambda) = \min_x \{f(x) + \mu g(x) + \lambda h(x)\}.$$

La ventaja del dual es que la definición de la función dual es un problema no restringido, y al mimos tiepo la restriccion misma para el problema dual es mucho mas simple (en particular lineal) que en el problema primario.

La idea del metodo dual consiste en aproximar la solución óptima del dual y entonces calcular una aproximación de la solución óptima del primario. Por conveniencia notacional, escribiremos

$$L(x, \mu, \lambda) = f(x) + \mu g(x) + \lambda h(x)$$

para el Lagrangiano del problema. En el siguiente algoritmo hemos usado el metodo del decenso mas rápido para la solución del dual. Recuerda que (Capítulo 3)

$$\nabla_\mu \theta(\mu, \lambda) = g(x), \quad \nabla_\lambda \theta(\mu, \lambda) = h(x)$$

si x *es tal que* $\theta(\mu, \lambda) = L(x, \mu, \lambda)$*.*

1. ***Inicialización.*** (μ_1, λ_1) *aproximación inicial con* $\mu_1 > 0$
2. ***Solución aproximada del primario.*** *Para* (μ_j, λ_j) *encuentre (aproxime) una solución óptima* x_j *para la función dual. Use un algoritmo no restringido para esto.*
3. ***Criterio de parada.****Para un valor umbral predeterminado* $\epsilon > 0$*, si*

   $$|h(x_j)| < \epsilon \quad |\mu_j g(x_j)| < \epsilon,$$

   detengase. La aproximación actual x_j *es lo suficientemente buena. De lo contrario continúe.*
4. ***Dirección de busqueda*** *Tome*

   $$d_k = \begin{cases} g_k(x_j), & si\ \mu_j^{(k)} > 0 \\ \max\{0, g_k(x_j)\}, & si\ \mu_j^{(k)} = 0. \end{cases}$$

donde $\mu^{(k)}$ representa la k-esima componente de μ, y

$$e_k = h_k(x_j)$$

5. ***Nueva aproximación.*** *Defina*

$$\mu_{j+1} = \mu_j + s_j d, \quad \lambda_{j+1} = \lambda_j + s_j e,$$

donde s_j se escoje para maximizar la función

$$\varphi(s) = \theta(\mu_j + sd, \lambda_j + se),$$

teniendo en cuenta que s esta restringido porque $\mu_j + sd \geq 0$. Regrese al paso 2

Metodo Lagrangiano aumentado. *Otra idea fuctifera consiste en usar la información proveniente de las condiciones de optimalidad. Para motivar la forma final del algoritmo que presentamos, enfoquemos primero en un PPNL con restricciones de igualdad, del tipo*

$$\textit{Minimizar} \quad f(x) \quad \textit{bajo} \quad h(x) = 0.$$

Sabemos que las soluciones óptimas deben satisfacer

$$\nabla f(x) + \lambda h(x) = 0.$$

para algun multiplicador apropiado λ. Supongamos que tenemos una aproximación para λ, λ_j. El asunto es como podemos usar esta λ_j para encontrar una aproximación a la solución óptima x_j, y simultaneamente mejorar la aproximación del multiplicador λ_{j+1} para proceder de manera iterativa. Obviamente el problema de optimización inicial es equivalente a

$$\textit{Minimizar} \quad f(x) + \lambda_j h(x) \quad \textit{bajo} \quad h(x) = 0$$

donde hemos incluido de manera trivial el parametro λ_j. Para resolver este problema introducimos una penalización cuadratica como la menciona anteriormente, y trabajamos el problema de minimizar la función de costo

$$f(x) + \lambda_j h(x) + \frac{r_j}{2}|h(x)|^2$$

para algun parametro de penalización r_j. La condición de optimalidad para este problema es

$$\nabla f(x) + \lambda_j h(x) + r_j h(x) \nabla h(x) = 0.$$

Si ademas suponemos que la aproximación x_j es buena, esta debe estar cerca de la verdadera solución óptima x, asi que si comparamos las condiciones de optimalidad para ambos problemas,

$$\begin{aligned} \nabla f(x) + \lambda \nabla h(x) &= 0, \\ \nabla f(x_j) + \lambda_j \nabla h(x_j) + r_j h(x_j) \nabla h(x_j) &= 0, \end{aligned}$$

llegamos a la conclusión que

$$\lambda_j + r_j h(x_j) \approx \lambda.$$

Estas ideas heuristicas (que pueden formalizarse) nos llevan al siguiente algoritmo iterativo

1. ***Iniciallización.*** *Tome $\lambda_1, r_1 > 1$, $c > 1$, y una tolerancia $\epsilon > 0$.*
2. ***Nueva aproximación*** *Encuentre una aproximación x_j al problema de minimización no restringido cuya función costo es*

$$f(x) + \lambda_j h(x) + \frac{r_j}{2}|h(x)|^2.$$

3. ***Criterio de Parada.*** *Si*

$$|\nabla f(x_j) + \lambda_j \nabla h(x_j)| < \epsilon,$$

detengase. x_j es una aproximación suficientemente buena. De lo contrario continúe.

4. ***Actualización*** *Actualice los valores del multiplicador y del parametro de penalización tomando*

$$\lambda_{j+1} = \lambda_j + r_j h(x_j), \quad r_{j+1} = cr_j$$

Regrese al paso 2

Para un PPNL general donde estan presentes ambos tipos de restricciones, en la forma de igualdades y en la forma de desigualdades, se pueden utilizar varios trucos para reducir el problema a el caso de restricciones de igualdad, como introducir nuevas variables mudas. Pero se puede usar el siguiente algoritmo que tiene el espíritu de los anteriores. El nuevo PPNL es

$$\textit{Minimizar} \quad f(x)$$

bajo

$$g(x) \leq 0, \quad h(x) = 0.$$

Por conveniencia notacional introduciremos el Lagrangiano aumentado.

$$L(x, \mu, \lambda, r) = f(x) + \mu g(x) + \lambda h(x) + \frac{r}{2}\left(\sum_i \text{máx}\{0, g_i(x)\}^2 + \sum_k h_k(x)^2\right).$$

1. ***Inicialización.*** *Escoja $\mu_1 \geq 0$, $\lambda_1, r > 0$, $c > 1$, y tolerancia $\epsilon > 0$.*
2. ***Nueva aproximación.*** *Encuentre una aproximación x_j para el mínimo del Lagrangiano aumentado $L(x, \mu_j, \lambda_j, r_j)$. Recuerde que la funcion*

$$\text{máx}\{0, g_i(x)\}^2$$

es diferenciable, asi que podemos usar un algoritmo típico para un problema sin restricciones como el gradiente conjugado o uno cuasi-Newton.

3. ***Criterio de Parada.*** *Si*

$$|\nabla f(x_j) + \mu_j \nabla g(x) + \lambda_j h(x_j)| < \epsilon, \quad |\mu_j g(x_j)| < \epsilon, \quad |h(x_j)| < \epsilon,$$

detengase y tome como una buena aproximación a x_j.

4. ***Actualización.*** *Actualice los valores de los multiplicadores y el parametro de penalización con las formulas*

$$\lambda_{j+1} = \lambda_j + r_j h(x_j), \quad \mu_{j+1} = \text{máx}\{0, \mu_j + r_j g(x_j)\}, \quad r_{j+1} c r_j.$$

Regrese al paso 2.

4.6. Comentarios Finales

La tarea de buscar soluciones óptimas globales para problemas de optimización es un trabajo tremendamente complejo y sutil. Nuestros lectores pueden tener la (falsa) impresión que los metodos que hemos descrito a traves de este capítulo o aquellos que hemos omitido son suficientes para resolver cualquier problema real. Esto esta muy lejos de ser cierto. En realidad, los metodos de gradiente estudiados anteriormente encuentran, con bastante éxito los mínimos locales de los funcionales objetivo. Graficamente podemos decir que usando una primera aproximación tomada .ªd hoc"', el algoritmo nos llevara al mínimo local ubicado en el "'valle de atracción"'de nuestra escogencia inicial. Si cambiamos esta iteración inicial, y esta se halla en un valle diferente, nuestra aproximación final dará un mínimo local distinto. Lo que es peor es que lo mas probable es que ninguno de los dos sera el mínimo global que estamos buscando, pero incluso si alguno de ellos es realmente el mínimo global, no hay forma que podamos estar seguros de esto. Todo tipo de algoritmos heuristicos han sido desarrollados a traves de los aos para aproximar soluciones globales a problemas tecnologicos de gran dimensión. La mayoria de ellos incorporan un ingrediente aleatorio de algun tipo. Entre ellos podemos citar tecnicas de descomposición, algoritmos de montecarlo y sus variantes, retorcido simulado y sus variantes, y algoritmos geneticos.

Este parrafo esta enfocado a insistir en la importancia de la convexidad desde el punto de vista de la aproximación númerica de las soluciones óptimas. ¿Qúe importancia tiene que tanto la función objetivo como aquella en las constantes sea convexa cuando se buscan soluciones óptimas globales? La respuesta esta dada en el teorema *3.14*. Una función definida sobre un conjunto convexo puede tener a lo mucho un valle, asi que cualquier algoritmo que nos de aproximaciones para un mínimo local, nos lleva al mínimo global. Esta es esencialmente la unica situación en la cual podemos asegurar que los algortimos mencionados atrapan mínimos globales. De lo contrario, el problema de encontrar y aproximar soluciones globales óptimas no tiene una solución rigurosa.

Finalmente volvemos a enfatizar que muchos de los algoritmos presentados en este capítulohan sido implementados en software comercial (vease la introducción a este capítulo). Aquellos que necesitan resolver problemas de optimización

de manera regular (programación lineal y no lineal) encontraran conveniente usar estas herramientas computacionales y concentrarse en la tarea de modelamiento directamente relacionada con los problemas de optimización. Es mas, una implementación efentiva de los algoritmos requiere una buena cantidad de experimentación, dado que el ajuste de los parametros es un ingrediente escencial para el exito. Esto no quiere decir que escribir un programa para el metodo del gradiente conjugado en alguno de los lenguajes de computación corrientes no es un buen ejercicio. De hecho una vez los estudiantes han tenido tal experiencia apreciaran profundamente la importancia del trabajo hecho por los especialistas en optimización númerica. Ciertos ejercicios de practica se proponen a continuación.

4.7. Ejercicios

1. Usando alguno de los metodos de busqueda lineal, encuentre el mínimo para las sigientes funciones, comenzando en los puntos dados.

   a) $f(x) = x^4 - x^2 + x - 1$, $x_0 = 1$;

   b) $g(x) = x^{16} + 3x^{14} - x^7 - 3$, $x_0 = -1$;

   c)
   $$h(x) = \int_{-1}^{x} \frac{se^s + s - 1}{2e^2 + 3} ds, \quad x_0 = -1$$

2. De un argumento de por que minimizando
   $$F(x) = \int_0^x f(s) ds$$
   podemos encontrar algunas soluciones pra la ecuación (no lineal) $f(x) = 0$. Aplique esto para encontrar algunas soluciones de las ecuaciones
   $$x^7 + x^4 + 1 = 0, \quad \log x + 6x = 0.$$

3. Considere la función
   $$f(x) = \frac{1}{100}(x^6 - 30x^4 + 192x^2 + 7x^3)$$
   Buscamos encontrar el mínimo global de f sobre la recta real $\mathbb{R}$. Aplique uno de los metodos de busqueda lineal comenzando con

   a) $x_1 = 1$;

   b) $x_1 = -1{,}2$;

   c) $x_1 = 5$;

   d) $x_1 = 3$;

   e) $x_1 = -8$;

$f)$ $x_1 = -3$;

¿Cúales son sus conclusiones? ¿Esta seguro acerca de cúal es el mínimo global de f? ¿Se le ocurre alguna manera para estar seguro de cúal es el mínimo global en este caso?

4. Trate de aproximar diferentes mínimos para las funciones

$$f(x) = \frac{2 - \sin|x| + x^2}{2 + \sin x}, \quad g(x) = \sin\left(\frac{x^2}{2}\right) + \sqrt{|x+1|}$$

usando alguno de los algoritmos de busqueda lineal comenzando con diferentes aproximaciones iniciales.

5. El sistema lineal $Ax = b$ puede ser resuelto de manera númerica minimizando la función

$$f(x) = \frac{1}{2}|Ax - b|^2$$

Si el sistema es soluble, entonces el mínimo global debe anularse, si no lo hace entonces el punto mínimo, es en algun sentido, el mas cercano a ala solución. Utilice algun algoritmo de decenso mas rápido para resolver numéricamente los sistemas

$a)$

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$b)$

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 1 \\ 2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

$c)$

$$A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 2 & 2 \\ -1 & 2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

6. Cuando la matriz A es simetrica y definida positiva, el sistema lineal $Ax = b$ es la ecuación para los puntos criticos de la forma cuadratica asociada

$$Q(x) = \frac{1}{2}x^T Ax - x^T b.$$

Argumente porque sucede esto, y encuentre una solución aproximada al sistema linal minimizando la forma cuadratica en este caso.

$$A = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & 1/4 \\ -1 & 1/4 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

7. Encuentre el mínimo de la función

$$P(x, y) = x^2 + 2y^2 - 2x - 8y$$

y aproximelo usando el metodo del decenso mas rapido comenzando el punto $(0, 0)$

8. Aproxime el minimo de la función

$$P(x, y) = 2x^2 - 2xy + y^y + 2x - 2y$$

con el metodo del decenso mas rapido. ¿Que resultados obtiene? Aproxime el mismso problema con el metodo del gradiente conjugado y compare los resultados.

9. Lo mismo que en el ejercicio anterior pra la función

$$P(x, y) = \frac{1}{4}(x^4 - 4xy + y^4).$$

10. Examine y estudie los siguientes problemas de minimizaciñ con los siguientes puntos iniciales.

   a) $100(x^2 - y)^2 + (1 - x)^2$, $(-1{,}2, 1)$;
   b) $(x^2 - y)^2 + (1 - x)^2$, $(-2, -2)$;
   c) $(x^2 - y)^2 + 100(1 - x)^2$, $(2, -2)$;
   d) $100(x^3 - y)^2 + (1 - x)^2$, $(-1{,}2, 1)$.

11. Aprixime el problema de minimización dado por

$$f(x_1, x_2.x_3, x_4) = (x_1 + 10x_2)^2 + 5(x_3 - x^4)^2 + (x_2 - 2x_3)^4 + (10x_1 - x_4)^4,$$

con aproximación inicial $(3, -1, 0, 1)$.

12. Nos gustaria aproximar el valor mínimo de la función $f(x, y) = y - x^2$ sobre el círculo unitario $x^2 + y^2 = 1$.

   a) Encuentre tal mínimo de manera exacta
   b) Aproxime la solución usando funciones de penalización de la forma

$$f_n(x, y) = y - x^2 + n(x^2 + y^2 - 1)^n$$

   para valores crecientes de n. Compare sus resultados.

13. . Encuentre una estrategia plausible para aproximar la solución óptima de

$$\text{Minimizar} \quad \frac{1}{2}x^T Ax$$

sujeto a

$$bx = c$$

donde A es una matriz simetrica definida positiva. Aplique esto para resolver el caso dado por

$$A = \begin{pmatrix} 3 & 3 & 1 \\ 3 & 5 & 3 \\ 1 & 3 & 3 \end{pmatrix} \quad b = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}, \quad c = 1.$$

14. Igual que el ejercicio anterior, pero cambiendo la restricción de igualdad $bx = c$ a la desigualdad $bx \leq c$. ¿Como aproximaria un problema similar bajo una restricción cuadratica del tipo

$$x^T Bx + bx \leq c$$

donde B es simetrica semidefinida positiva? Aplique su metodo para resolver

$$\text{Minimizar} \quad |x|^2$$

bajo

$$(2x_1 - x_2)^2 + (x_3 - 2)^2 \leq 1.$$

15. Escoja algunos de los ejercicios propuestos en el capítulo 3 para un conjunto particular de datos y aproxime sus soluciones.

Capítulo 5

Problemas Varacionales y Programación Dinamica

5.1. Introduccion

Comenzamos este capítulo con el analísis de problemas de optimización de otra naturaleza. Especificamente, este capítulo esta dedicado a problemas variacionales de encontrar el infimo de las integrales

$$\int_{\Omega} F(x, u(x), \nabla u(x))dx,$$

donde $\Omega \subset \mathbb{R}^n$, las funciones $u : \Omega \to \mathbb{R}$ deben ser diferenciables, y estas deben normalmente estar restringidas de alguna manera como teniendo fijos sus valores en $\partial\Omega$ por alguna funcin u_0 asignada con anterioridad, i.e $u = u_0$ en $\partial\Omega$. El integrando (o Lagrangiano)

$$F : \omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$$

Caracteriza tal problema. La tarea propuesta consiste en encontrar una función U, admisible de acuerdo con las restricciones que hemos impuesto en las funciones competentes, de tal manera que la integral

$$\int_{\Omega} F(x, U(x), \nabla U(x))dx$$

es mas pequeña (o igual) que la misma integral para cualquier otra función factible u. Si usamos la notación

$$I(u) = \int_{\Omega} F(x, u(x), \nabla u(x))dx.$$

estamos interesados en entender el siguiente problema de optimización

$$\text{Minimizar} \quad I(u)$$

sujeto a las restricciones sobre las funciones u, por ejemplo

$$u(x) = u_0(x), \quad x \in \partial\Omega$$

Por lo tanto esto es un problema de optimización en el cual las funciones admisibles remplazan los vectores factibles.

En este capítulo introductorio estaremos interesados principalmente en problemas variacionales en una dimensión donde $\Omega = (a, b)$ es un intervalo en $\mathbb{R}$, y las funciones admisibles $u : (a, b) \to \mathbb{R}$ a menudo se les pide que satisfagan las condiciones de frontera

$$u(a) = A, \quad u(b) = B,$$

para valores conocidos A y B. En este caso

$$I(u) = \int_a^b F(x, u(x), u'(x))dx$$

donde ahora el integrando F es una función de tres (o menos) variables. Hemos mencionados algunos de estos ejemplos en el capítulo 1, y hemos tratado de convencer al lector que resolver este tipo de problemas o al menos aproximar las soluciones óptimas apropiadamente) pude ser importante. En este capítulo estudiaremos y resolveremos mas ejemplos, y aprenderemos las tecnicas principales para lidiar con tales problemas.

Hay una gran variedad de problemas variacionales. El ingrediente en comun es que los costos estan tipicamnente representados por una integral como la anterior. Pero las restricciones adicionales pueden variar de ejemplo a ejemplo. Podemos clasificar estas restricciones como sigue.

1. **Condiciones de Frontera.** Una de las situaciones mas comunes corresponde a haber preescrito valores en toda la frontera $\partial\Omega$, pero otras posibilidades incluyen esta restricción solo en parte de la frontera (en particular cuando no hay ninguna condicion en absoluto) y no preescribir los valores en la frontera sino aquellos de su derivada, o algunas de sus derivadas.

2. **Restricciones Integrales.** Estas requieren que las funciones admisibles cumplan con restricciones del tipo

   $$\int_\Omega G(x, u(x), \nabla u(x))dx = \alpha$$

   donde

   $$G : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^d, \quad \alpha \in \mathbb{R},$$

   son conocidas, algunas de estas constantes pueden venir en forma de desigualdades.

3. **Restricciones Puntuales.** Establecen que las funciones factibles deben respetar la condición

   $$G(x, u(x), \nabla u(x)) = 0$$

para todo x en Ω, donde G es una función conocida como antes. Tambien podriamos tener algunas desigualdades

Finalmente es importante anotar que algunas de las tecnicas que seran discutidas puden ser extendidas sin mayor cambio a situaciones en las cuales los funcionales de costo incluyen una dependencia en derivadas de orden superior (o no tienen tal dependencia). Trabajaremos con algunas de estas situaciones.

Otro aspecto bastante cercano a los problemas variacionales es la programación dinamica. En el caso discreto, discutiremos brevemente el principio de fondo que lleva a soluciones óptimas. En el caso continuo, estableceremos heuristicamente la ecuación de Bellman de programación dinamica, la cual nos ayudara a deducir el principio del máximo de Pontryagin en el capítulo siguiente.

Es justo enfatizar la importancia de los problemas variacionales en varios campos de la ciencia y la ingenieria. Dificilmente podemos listarlos todos ellos aqui: mecanica, elasticidad (lineal y no lineal), medios continuos, dinamica, comportamiento de materiales, fluidos, etc. Algunos de nuestros ejemplos ilustraran de una manera simple y directa la relevancia y el rol jugador por las formulaciones y tecnicas variacionales. Esto no es de sorprender, dado que la naturaleza, asi como el hombre de alguna manera u otra siempre busca lo mejor.

5.2. La ecuación de Euler-Lagrange: Ejemplos

La ecuación de Euler Lagrange (E-L) asociada con un problema variacional juega el mismo rol que las condiciones necesarias de optimalidad en un problema de programación (las condiciones KKT). Como ahora sabemos, tales condiciones necesarias para optimalidad nos dan restricciones que las soluciones óptimas (entre otros vectores factibles) deben cumplir. Explotando estas condiciones, se pueden encontrar o aproximar soluciones óptimas en una variedad de situaciones. Tambien hemos enfatizado el papel central de la noción de convexidad en asegurar que las condiciones de optimalidad son de hecho suficientes para las soluciones óptimas. Por lo tanto no es de sorprender que la convexidad tambien jugara un papel importante en esto. La convexidad siempre es el concepto clave en los problemas de minimización de cualquier tipo.

En general un problema variacional esta caracterizado por un costo funcional integral

$$I(u) = \int_{\Omega} F(x, u(x), \nabla u(x))dx,$$

donde $\Omega \subset \mathbb{R}^n, u : \Omega \to \mathbb{R}$ es diferenciable, y el integrando del costo

$$F(x, \lambda, \xi) : \Omega \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$$

se supone diferenciable, de hecho dos veces diferenciable con respecto a las variables (λ, ξ). Note que

$$\xi = (\xi_1, \xi_2, \ldots, \xi_N).$$

Nos concentraremos en la situación en la cual ademas tenemos restricciones adicionales de factibilidad del tipo

$$u(x) = u_o(x), \quad x \in \partial\Omega,$$

dado que esta es la situación mas típica. Mas adelante cuando nos restrinjamos al caso $N = 1$ consideraremos otros casos.

El siguiente resultado es la pista para encontrar soluciones óptimas para este tipo de problema.

Teorema 5.1 *(Ecuación de Euler-Lagrange) Bajo las hipótesis anteriores:*

1. *Si u es una solución óptima, entonces u debe tambien ser una solución del problema (E-L)*

$$\begin{aligned} \operatorname{div}(F_\xi(x, u(x), \nabla u(x))) = F_\lambda(x, u(x) m \nabla u(x)) \quad in \quad \Omega \\ u = u_0 \quad en \quad \partial\Omega \end{aligned}$$

2. *Si u satisface E-L y F es convexa con respecto a las variables (λ, ξ) para cada $x \in \Omega$ fijo, entonces u tambien es una solución óptima para el problema variacional.*

3. *Si ademas F es estrictamente convexo con respecto a (λ, ξ) para cada $x \in \Omega$, la solución óptima si existe es única.*

Antes de explicar de donde viene esta ecuación E-L. Es importante estar convencido de su relevancia y aplicabilidad para resolver algunos problemas variacionales. Estudiaremos algunas situaciones particulares, incluyendo algunos de los casos del primer capítulo. La mayoria de estos casos corresponden al caso uno dimensional, cuando $N = 1$ y Ω es de hecho un intervalo abierto (a, b) en la recta real. La ecuación E-L es en estecaso una ecuación diferencial ordinaria de segundo orden completada con los valores de frontera apropiados

$$\begin{aligned} \frac{d}{dx}[F_\xi(x, u(x), u'(x)))] = F\lambda(x, u(x), u'(x)), \quad x \in (a, b), \\ u(a) = A, \quad u(b) = B \end{aligned}$$

donde normalmente A y B estan dados. En esta situación unidimensional cubriremos varias posiblidades para ganar familiaridad con la ecuación E-L.

1. Para empezar, supongamos la situación mas simple, en la cual F depende exclusivamente de ξ. En este caso $F = F(\xi)$, y la ecuación E-L se simplifica a

$$\frac{d}{dx}[F'(u'(x))] = 0,$$

lo que implica que

$$F'(u'(x)) = k,$$

una constante. Evidentemente, este ultimo requerimiento se cumple si tomamos u'constante en el intervalo (a,b), y esto se cumple si u es la linea recta que une los puntos (a, A) y (b, B), tal función lineal (afín) es siempre una solución de E-L su F solo depende del valor de la variable derivada ξ. Si ademas F es convexa esta función es un minimizador del problma. Y si la función es F es estrictamente convexa, esta función lineal es el unico minimizador del problema. En caso que F no sea convexa, incluso cuando la función lineal sea una solución de E-L, puede no ser un minimizador como muestra el siguiente ejemplo.

Ejemplo 5.2 *(Ejemplo no convexo) Tomemos*

$$F(\xi) = e^{\xi^2}, \quad a = A = B = 0, \quad b = 1.$$

En esta situación la función lineal que pasa por los puntos $(0,0)$,$(1,0)$ *es la función* u_0 *que se anula en todo el intervalo* $(0,1)$. *Su costo es* 1. *Sin embargo aseguramos que el infimo de las integrales*

$$I(u) = \int_0^1 e^{-u'(x)^2} dx$$

suejeto a

$$u(0) = u(1) = 0$$

es 0. Para mostrarlo considere la siguiente sequencia de funciones factibles

$$u_j(x) = j\left(x - \frac{1}{2}\right)^2 - \frac{j}{4}$$

No es deficil confirmar que $I(u_j) \searrow 0$, *asi que el infimo anterior se anula. Por lo tanto* u_0 *es una solución del problema* $E - L$ *asociado, pero no es un minimizador. Lo que falla en esta situación es la convexidad de* F. *Es mas, no existe minimizador para este problema. ya que tal función tendria que satisfacer*

$$\int_0^1 e^{-v'(x)^2} dx = 0$$

y esto es imposible.

Ejemplo 5.3 *(Geodesicas en un cilindro) Sea* C *el cilindro dado por la ecuación*

$$x^2 + y^2 = 1,$$

y sean P *y* Q *dos puntos distintos en* C. *Nos gustaria encontrar el camino mas corto sobre* C *para ir de* P *a* Q. *Sin perdida de generalidad podemos suponer que* $P = (1,0,0)$. *Naturalmente pondremos este problema en coordenadas cilindricas* (r, θ, z) *con*

$$r = 1, \quad -\pi \leq \theta \leq \pi$$

definiendo a C. Sea $Q = (1, \theta_0, z_0)$ (vea la figura 5.1).

Por consideraciones de simetria, es suficiente tratar el caso $0 < \theta_0 \leq \pi$ (¿cúal es la geodesica cuando $\theta_0 = 0$?). Podemos representar una curva arbitraria que una $(1, 0, 0)$ y $(1, \theta_0, z_0)$ en la forma (en coordenadas cartesiamas)

$$\sigma(\theta) = (\cos\theta, \sin\theta, z(\theta)), \quad \theta \in (0, \theta_0),$$

dado que las geodesicas cortan cada linea vertical en C a lo mas una sola vez (¿por qué?) y de esta manera las primeras dos componentes de σ se puden considerar trigonometricas. De esta manera cualquier tal curva esta completamente determinada por la función $z(\theta)$. Tambien debemos exigir

$$z(0) = 0 \quad z(\theta_0) = z_0.$$

El funcional de costo que debemos minimizar es aquel que representa la longitud de σ:

$$I(z) = \int_0^{\theta_0} \sqrt{1 + z'(\theta^2)} d\theta$$

Sabemos que la función

$$F(\xi) = \sqrt{1 + \xi^2}$$

es estrictamente convexa (Capítulo 3), y solo depende de la variable derivada. Por lo discutido anteriormente la función lineal

$$z(\theta) = \frac{z_0}{\theta_0}\theta$$

representa la unica geodesica que une estos dos puntos. Esta función lineal en coordenadas cilindricas representa un arco de una helice sobre C.

2. Cuando el integrando F depende de ambos x y ξ, E-L se converte en

$$\frac{d}{dx}[F_\xi(x, u'(x))] = 0,$$

o de manera equivalente

$$F_\xi(x, u'(x)) = k, \quad \text{(constante)}$$

Dependiendo de la forma particular de F esta ecuación puede ser resuelta de manera analitica o no.

Ejemplo 5.4 *(Ejemplo de Weiertrass) Sea*

$$F(x, \xi) = x\xi^2, \quad x \in (0, 1).$$

y $u(0) = 1$, $u(1) = 0$ en los puntos extremos del intervalo. En este ejemplo, E-L es

$$xu'(x) = x, \quad u(x) = c\log x + d$$

Curriosamente esta familia de soluciones es incapaz de cumplir la condición de frontera en el extremo izquierdo $u(0) = 1$, ya que $u(0)$ no esta definido para ninguna de las soluciónes. Por lo tanto el problema variacional puede no tener soluciones óptimas. Este es el caso, ya que la familia de funciones

$$u_j(x) = \begin{cases} 1 & x \in (0, 1/j), \\ -\log x/\log j, & x \in (1/j, 1), \end{cases}$$

es minimizante para

$$I(u) = \int_0^1 xu'(x)^2 dx, \quad u(0) = 1, \quad u(1) = 0,$$

En el sentido que $I(u_j) \searrow 0$., Ademas para cualquier función u, tenemos $I(u) > 0$, y por lo tanto no hay solución óptima para este problema variacional.

Ejemplo 5.5 *(La braquistocrona) Uno de los problemas variacionales mas famoso de todos los tiempos el es de la braquistocrona. Coloquemos el eje x en direccion vertical paralela a la acción de la gravedad, y el eje y perpendicular a este. Suponga que tenemos dos puntos P y Q a diferentes alturas. Sin perdida de generalidad podemos tomar P en el origen y $Q = (a, A)$ con ambas a, A positivas. La tarea consiste en encontrar el camino entre P y Q de tal manera que una unidad de masa gasta el menor tiempo posible moviendose de Q a P unicamente bajo la acción de la gravedad (sin fricción).*

Evidentemente el camino óptimo puede ser representado en la forma $y(x)$ para alguna función y a determinar. Lo que queremos decir con esto es que los caminos que no son monotonos (con saltos) obviamente tardaran mas que aquellos monotonos. Asuma que $y(x)$ es tal camino, asi que $y(0) = 0$, $y(a) = A$.

El tiempo de transito para tal camino esta dado por la integral

$$\int_0^a \frac{ds}{v},$$

donde ds es el elemento diferencial de longitud de arco dado por

$$ds = \sqrt{1 + y'(x)^2}dx,$$

y v es la velocidad por la gravedad a una altura x. De acuerdo a una formula

$$v = \sqrt{2gx}$$

Estamos interesados en encontrar la curva $y(x)$, $0 \leq x \leq a$, que minimiza la integral de transito (ignorando constantes positivas que no interfiren con la optimización)

$$I(y) = \int_0^a \sqrt{\frac{1 + y'(x)^2}{x}}dx,$$

y $y(0) = 0$, $y(a) = A$. *Advertimos al lector que la solución no es ni una linea recta, ni un circulo. De hecho en el caso particular* $a = A = 1$, *¿Pueden nuestros lectores decidir cual curva resulta en menos tiempo de caida? ¿la linea* $y = x$, *o el circulo* $y = 1 - \sqrt{1 - x^2}$*? En nuestro problema la función integrando* F *es*

$$F(x, \xi) = \frac{\sqrt{1 + \xi^2}}{\sqrt{x}}.$$

Que es una funcion estrictamente convexa con respecto a ξ, *asi que ignorando la dificultad cuando* $x = 0$, *las soluciones óptimas se deben buscar examinando la ecuación E-L. En este caso debemos resolver*

$$\frac{y'}{\sqrt{x}\sqrt{1 + (y')^2}} = \frac{1}{c}, \quad \frac{(y')^2}{1 + (y')^2} = \frac{x}{c^2}.$$

Esto nos lleva a

$$y'(x)^2 = \frac{x}{c^2 - x}, \quad y(x) = \int_0^x \sqrt{\frac{s}{c^2 - s}} ds,$$

Donde la constante c se determinara de manera que

$$A = \int_0^a \sqrt{\frac{s}{c^2 - s}} ds.$$

¿Puede argumentar el lector porque hemos escogido el signo positivo en la raiz cuadrada anterior?

Para encontrar una forma mas explicita de la solución, cambiaremos las variables en la integral y de la siguiente manera.

$$s(r) = \frac{c^2}{2}(1 - \cos r) = c^2 \sin^2(r/2).$$

Entonces

$$y(t) = c^2 \int_0^t \sin^2(r/2) dr = \frac{c^2}{2}(t - \sin t),$$

donde

$$x(t) = \frac{c^2}{2}(1 - \cos t) = c^2 \sin^2(t/2).$$

En forma parametrica,

$$(x(t), y(t)) = (C(1 - \cos t), C(t - \sin t)), \quad 0 \leq t \leq t_0.$$

Esta solución ya satisface $x(0) = y(0) = 0$. *las constantes* C *y* t_0 *deben ser encontradas con las condiciones* $x(t_0) = a$, $y(t_0) = A$. *Esta curva es el arco de una cicloide (Figura 5.3).*

3. Finalmente analizaremos el caso en el cual $F = F(\lambda, \xi)$. En este caso la ecuación E-L tiene la forma

$$\frac{d}{dx}[F_\xi(u(x), u'(x))] = F_\lambda(u(x), u'(x)).$$

Es cuestion de simple aritmetica ver que este ecuación (en el caso que $F = F(\lambda, \xi)$ y solo en ese caso) puede ser reescrita como

$$\frac{d}{dx}[F(u(x), u'(x)) - u'(x)F_\xi(u(x), u'(x))] = 0$$

En algunas situaciones este forma de la ecuación puede ser mas apropiada cuando se estan buscando soluciones. Pero en otras ocasiones esto puede no suceder.

Ejemplo 5.6 *(Superficies Mínimas de Revolución) Nos gustaria identificar funciones $u(x)$ definidas en el intervalo (a, b) de tal manera que $u(a) = A$, $u(b) = B$, y cuya grafica genera por revolucion alrededor del eje X, la superficie de revolución con menor área (Figura 5.4)*

Sabemos del cálculo que esta area esta dada, exepto por un factor positivo constante, por la integral.

$$I(u) = \int_a^b u(x)\sqrt{1 + u'(x)}dx.$$

Buscamos la función que minimice esta integral entre todas aquellas que satisfacen las condiciones en ambos extremos del intervalo. El integrando para este ejemplo es

$$F(\lambda, \xi) = \lambda\sqrt{1 + \xi^2}$$

Esta función es convexa en ξ para un λ fijo dado que $\lambda \geq 0$. Es incluso estrictamente convexa en ξ si lambda > 0. Sin embargo no es conjuntamente convexa en (λ, ξ) (¿Por qu?). Por lo tanto en principio no podemos usar el Teorema 5.1. Sin embargo, para que el resultado sea cierto solo se necestita convexidad con respecto a ξ para (x, λ) fijos. La demostración de esto esta por fuera del objetivo de este texto, pero es importante tener esto en mente para el ejemplo siguiente.

Teorema 5.7 *([19]) Si u es la unica solución de E-L y F es convexo con respecto a la variable ξ para cada $x \in \Omega$ fijo y $\lambda \in \mathbb{R}$, entonces u es tambien una solución óptima para el problema variacional.*

Por lo tanto podemos proceder con el estudio de la ecuación E-L para buscar soluciones óptimas

La segunda forma de la ecuación E-L para este ejemplo es

$$u\sqrt{1 + (u')^2} - u'\frac{uu'}{\sqrt{1 + (u')^2}} = c$$

Despues de varias manipulaciones, separando varialbes, llegamos a

$$\frac{du}{\sqrt{u^2 - c^2}} = \frac{dx}{c}$$

En este caso particular no importa el signo que tomemos para la raiz cuadrada. Invitamos a nuestros lectores que comprueben esto. Un cambio de variables usando funciones hiperbolicas nos lleva a la forma final de la solución

$$u(x) = c \cosh\left(\frac{x}{c} + d\right),$$

la cual es una catenaria. Esta solución es única.

Al ajustar los valores en los puntos extremos encontramos algunas dificultades. Suponga, por ejemplo que exigimos que $u(0) = 1$, $u(b) = B$. *La primera condición implica que*

$$c = \frac{1}{\cosh d}$$

mientras que la segunda implica

$$B = \frac{\cosh(b \cosh d + d)}{\cosh d}$$

Esta condicion no se puede satisfacer siempre. Teniendo en cuenta que

$$\cosh t > |t|$$

para todo real t *resulta que*

$$B > \frac{b \cosh d + d}{\cosh d} \geq \frac{b \cosh d - |d|}{\cosh d} = b - \frac{|d|}{\cosh d} \geq b - 1$$

Esta cedena de desigualdades implica que si al principio hubieramos impuesto $B \leq b - 1$, *no habria forma de ajustar el valor en el extremo derecho, y por lo tanto el problema puede no tener soluciones óptimas.*

Esta observacion tambien puede ser ilustrada fisicamente de una forma bastante atractiva. La forma adoptada por una burbuja de jabon que se adiherea un anillo (no planar) esta realacionado con la tensión superficial de tal manera que la forma adoptada por la pelicula de jabon sera la que minimice la tensión superficial. Esta cantidad es proporcional al área generada porla superficia, asi que determinar la forma óptima es equivalentea encontrar la superficie de mínima área. En el caso de superficies de revolución, como aquellas que estamos considerando en este ejemplo, podemos imaginar dos anillos concentricos de distintos radio, y a una distancia pequeña. Una pelicula de jabon pegada a los dos anillos mostrara la forma dada por la catenaria que encontramos anteriormente. Pero a medida que alejemos los anillos uno del otro la pelicula se estirara hasta cierto punto donde la pelicula se degenerara a dos circulos en ambos anillos. Esta transición esta reflejada en nuestros calculos.

La explicación de este hecho depende de nuestro entendimiento de nuestro problema variacional. Cuando los dos anillos estan cerca, la catenaria da la superficie mínima de revolución con área menor que aquella de los dos circulos que definen los anillos. A medida que alejamos los anillo, el área de la catenaria crece hasta el punto donde iguala aquella de los dos circulos. Si la alejamos mas, ninguna superficie de revolución tendra menor área que los dos circulos, y por lo tanto la pelicula se separa. Esta solución óptima (los dos discos) no puede ser descrita como una superficie de revolución de alguna función u asi que en realidad nuestro problema variacional no tiene una solución óptima. Una formulación mas general del problema requeriria incorporal estas funciones especiales. Sin embargo es cierto que podemos aproximar esta solución especial tanto como querramos mediante una sucesión de superficies de revolución admisibles u_j de acuerdo co la figura 5.5

4. Finalmente examiramos varias situaciones sencillas en varias dimensiones.

Ejemplo 5.8 *(Integral de Dirichlet) Sea $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ una función que satisface $u = u_0$ en $\partial\Omega$ donde u_0 es una función fija. Nos gustaria identificar la función u que minimice la integral*

$$I(u) = \frac{1}{2}\int_\Omega |\nabla u(x)|^2 dx.$$

Podemos imaginar situaciones en las cuales $N = 2$ o $N = 3$, y Ω es un circulo en el plano o una bola en el espacio. Es por lo tanto un problema de calculo variacional. Notese que el integrando

$$F(\xi) = \frac{1}{2}|\xi|^2$$

es estrictamente convexo, asi que existe una solución óptima, y es unica (vease Teorema 5.1) Tal función debe ser la solución de la ecuación E-L. En este caso la ecuación es

$$\operatorname{div}(F_\xi(\nabla u(x))) = 0$$

i.e

$$\Delta u = 0$$

Por lo tanto la función que minimiza el cuadrado de la norma del gradiente es la función armonica que respete las condiciones de frontera en $\partial\Omega$. Normalmente esto se interpreta diciendo que la función armonica corresponde a un estado de equilibrio estable con respecto a una energia proporcional al cuadrado del gradiente.

Ejemplo 5.9 *(Ecuación de onda) Para el caso particular en el cual*

$$F(\xi_1, \xi_2) = \frac{1}{2}(\xi_1^2 - \xi_2^2),$$

la ecuación E-L es precisamente la eciación de onda

$$u_{xx} - u_{yy} = 0.$$

Sin embargo en este caso F no es convexo, asi que no podemos hablar de minimización. A pesar de esto para este tipo de ecuación existe una amplia teoria de principios variacionales en mecanica donde se postula el principio de Hamilton de mínima acción y el papel central se mueve del concepto de minimizador a aquel de estado estacionario.

Un ejemplo simple puede ayudarnos a entender un poco mejor lo que significan las lineas anteriores. Suponga que una particula de masa m viaja en linea recta bajo la acción de un campo $F(t, x)$ que depende de la posición y del tiempo. Si $u(t)$ indica la posición de la particula en un tiempo t. La Ley de Newton nos dice que

$$\frac{d}{dt}(mu'(t)) = F(t, u(t)).$$

La pregunta es si podemos crear una función $L(t, \lambda, \xi)$ de tal manera que la ecuación E-L asociada con el funcional

$$I(u)) \int_0^{t_0} L(t, \lambda, \xi)dt$$

resulta ser exactamente la ley de Newton. En esta situación simplificada no es dificil. Si suponemos que el campo F es conservativo con potencial $U(t, \lambda)$ de tal manera que $U_\lambda = -F$, entonces podemos tomar

$$L(t, \lambda, \xi) = \frac{1}{2}\xi^2 - U(t, \lambda).$$

Esta función es convexa en ξ. Por el teorema 5.7, nos gustaria concluir que el movimiento de la particula toma lugar de acuerdo al principio de energia mínima (mínima acción) medido por I. La función L se llama el Lagrangiano, y al funcional I se le llama la integral de acción. Un estudio copleto de estos temas excede los objetivos de este libro.

Ejemplo 5.10 *(Superficies Mínimas) Consideramos el problema de superficies minimas, no necesariamente de recolución alrededor de un eje. Dada una región del plano Ω y una función u_0 sobre $\partial\Omega$ (el borde), el area de la grafica de una función u esta dada por dada por la integral*

$$I(u) = \int_\Omega \sqrt{1 + |\nabla u(x)|^2}dx.$$

Estamos buscando una función u que minimice esta integral entre todas aquellas funciones que tienen los mismos valores u_0 en $\partial\Omega$. Dado que la función

$$F(\xi) = \sqrt{1 + |\xi|^2}$$

es estrictamente convexa, si existe una solución esta debe ser unica. La ecuación E-L es

$$\operatorname{div}\left(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\right)=0.$$

Esta ecuación diferencial parcial es una ecuación bastante complicada por razones que van mas alla del objetivo de este libro. En el caso $N=2$, la ecuación puede ser reescrita como

$$(1+u_y^2)u_{xx}-2u_xu_yu_{xy}+(1+u_x^2)u_{yy}=0.$$

5.3. La ecuación de Euler-Lagrange: Justificación

Despues de convencernos por numerosos ejemplos de la importancia de la ecuación E-L para encontrar soluciones óptimas para problemas variacionales, vale la pena explorar los origenes de esta ecuación y porque las soluciones óptimas deben tambien ser soluciones de esta ecuación. Estudiaremos estos temas entendiendo la idea de fondo y omitiendo varios detalles tecnicos que son irrelevantes a nuestra discución. Veremos primero el caso unidimensional para entender mejor los origenes y luego indicaremos los cambios para el caso multidimensional.

Teorema 5.11 *(Ecuación de Euler-Lagrange) Bajo las hipotesis mencionadas anteriormente:*

1. *Si u es una solución óptima, entonces u tambien debe ser una solución del problema E-L*

$$\begin{aligned}\operatorname{div}(F_\xi(x,u(x),\nabla u(x)))=F_\lambda(x,u(x),\nabla u(x)) \quad &in \quad \Omega,\\ u=u_0 \quad &en \quad \partial\Omega\end{aligned}$$

2. *Si u satisface E-L y F es convexa con respecto a las variables (λ,ξ) para cada x fijo en Ω, entonces u tambien es una solución óptima del problema variacional.*

3. *Si ademas F es estricamente convexo con respecto a (λ,ξ) para cada $x\in\Omega$, la solución óptima u si existe es única.*

Para el caso unidimensional esta ecuación se simplifica a

$$\frac{d}{dx}[F_\xi(x,u(x),u'(x))]=F_\lambda(x,u(x),u'(x)),$$

junto con las condiciones de frontera $u(a)=A$, $u(b)=B$, donde

$$F=F(x,\lambda,\xi),\quad I(u)=\int_a^b F(x,u(x),u'(x))dx.$$

sea φ una función fija que satisface el requisito $\varphi(a) = \varphi(b) = 0$, y consideremos la función de una sola variable.

$$g(t) = I(u + t\varphi) = \int_a^b F(x, u(x) + t\varphi(x), u'(x) + t\varphi'(x))dx,$$

donde asumimos que u es una solución óptima que nos da el mínimo valor de las integrales entre todas las funciones factibles. Para cada valor de $t \in \mathbb{R}$, la función $u + t\varphi$ es una función admisible ya que φ se anula en los extremos del intervalo (a, b). Por lo tanto g tiene un mínimo (global) para $t = 0$. Una condición necesaria para que tal mínimo ocurra es que la derivada se anule en tal punto. Si derivamos bajo el simbolo de integral en la definición de g y evaluamos en $t = 0$ obtenemos

$$0 = \int_a^b [F_\lambda(x, u(x), u'(x))\varphi(x) + F_\xi(x, u(x), u'(x))\varphi'(x)]dx.$$

Si integramos por partes el segundo termino, teniendo en cuenta que $\varphi(a) = \varphi(b) = 0$, tenemos

$$0 = \int_a^b \left[F_\lambda(x, u(x), u'(x)) - \frac{d}{dx}F_\xi(x, u(x), u'(x))\right] \varphi(x)dx.$$

Dado que φ es arbitrario, excepto por sus valores en la frontera, la identidad anterior solo puede ocurrir si la expresión entre parentesis cuadrados es identicamente cero: La ecuación E-L. Esta es la primera parte del teorema. Supongamos ahora que el integrando $F(x, \lambda, \xi)$ es conjuntamente convexo en las variables (λ, ξ) para cada x fijo en (a, b), y que u es una solución de E-L junto con las condiciones de frontera apropiadas. Sea v cualquier otra función factible tal que $v(a) = A$, $v(b) = B$. Por la convexidad de F (Capítulo 3),

$$I(v) - I(u) = \int_a^b [F(x, v(x), v'(x)) - F(x, u(x), u'(x))]dx$$

$$\geq \int_a^b [F_\lambda(x, u(x), u'(x))(v(x) - u(x)) + F_\xi(x, u(x), u'(x))(v(x) - u(x))dx.$$

Si integramos por partes el segundo termino como antes, notamos que obtenemos exactamente la ecuación E-L para u, asi que si u es una solución la conclusión es

$$I(v) - I(u) \geq 0,$$

y u es realmente una solución óptima para el problema. Esto demuestra la suficiencia.

Finalmente nos gustaria provar la unicidad de las soluciones óptimas bajo la convexidad estricta de F. La forma mas facil de manejar la convexidad estricta en este contexto consiste en exigir que la igualdad

$$f\left(\frac{1}{2}x + \frac{1}{2}y\right) = \frac{1}{2}f(x) + \frac{1}{2}f(y)$$

automaticamente implica $x = y$ si f es una función estrictamente convexa. Tratemos de llegar a tal situación cuando asumimos que $F(x, \cdot, \cdot)$ es estrictamente convexa.

Suponga que nuestro problema variacional admite dos soluciones óptimas u, v. Debido a la convexidad no es dificil deducir

$$\frac{1}{2}I(u) + \frac{1}{2}I(v) - I(\frac{1}{2}u + \frac{1}{2}v) \geq 0$$

Si denotamos por m el valor del mínimo i.e. $I(u) = I(v) = m$, entonces

$$m - I(\frac{1}{2}u + \frac{1}{2}v) \geq 0$$

Pero por otro lado, dado que m es el valor mínimo

$$I(\frac{1}{2}u + \frac{1}{2}v) \geq m$$

De esto podemos concluir

$$I(\frac{1}{2}u + \frac{1}{2}v) = m$$

y dado que

$$\begin{aligned} 0 = \int_a^b \Bigg[& \frac{1}{2}F(x, u(x), u'(x)) + \frac{1}{2}F(x, v(x), v'(x)) \\ & -F\left(x, \frac{1}{2}u(x) + \frac{1}{2}v(x), \frac{1}{2}u'(x) + \frac{1}{2}v'(x)\right)\Bigg] dx. \end{aligned}$$

Pero el integrando anterior es no negativo, por la convexidad de F. La unica posiblidad para una función no negativa cuya integral se anula es ser la función nula, asi que

$$\begin{aligned} & \frac{1}{2}F(x, u(x), u'(x)) + \frac{1}{2}F(x, v(x), v'(x)) \\ & -F\left(x, \frac{1}{2}u(x) + \frac{1}{2}v(x), \frac{1}{2}u'(x) + \frac{1}{2}v'(x)\right) = 0. \end{aligned}$$

Por lo mencionado anteriormente esto implica que $u = v$, y que la solución óptima si existe es única.

Para el caso de un problema en varias variables, el argumento es formalmente el mismo. Los cambis estan en la forma en que se hace la integración por partes dado el teorema de la divergencia. Por lo tanto si φ se anula en $\partial\Omega$, de manera que las contribuciones de la frontera son cero, tendremos

$$\begin{aligned} 0 &= \int_\Omega [F_\lambda(x, u(x)\nabla u(x))\varphi(x) + F_\xi(x, u(x)\nabla u(x))\nabla\varphi(x)]dx \\ &= \int_\Omega [F_\lambda(x, u(x)\nabla u(x)) - \operatorname{div}(F_\xi(x, u(x)\nabla u(x)))]\varphi(x)dx. \end{aligned}$$

Y se sigue la prueba de la misma manera

5.4. Condiciones Naturales de Frontera

Es impotante profundizar en la manera como deducimos E-L. Para enfatizar esto, veremos una situación típica en la cual el valor en uno de los puntos extremos esta libre en un problema variacional de dimensión uno. Nos gustaria encontrar el mínimo del funcional

$$I(u) = \int_a^b F(x, u(x), u'(x))dx,$$

donde les exigimos a las funciones u solamente que satisfagan $u(a) = A$, pero no se les exige nada en el punto extremo derecho, asi que el conjunto de funciones factibles es mas grande comparado con la situación en la cual fijamos el valor en ese punto extremo. Sospechamos que E-L junto con $u(a) = A$ puede no ser sufiente para determinar completamente la solución óptima u. De alguna manera, la condición de tener un punto extremo libre debe imponer alguna condición adicional en las soluciones óptimas. Nuestras sospechas son ciertas. Si regresamos a la deducción de la ecuación E-L, notamos que el paso en el cual los valores en los extremos eran importantes era en la escogencia de la función auxiliar φ, una función que debe anularse en a y en b. Si ahora dejamos el valor en b libre, equivale a considerar φ arbitraria excepto por $\varphi(a) = 0$, sin exigirle nada en el punto b. Esta información se uso durante la integración por partes. Si no tenemos $\varphi(b) = 0$ obtendriamos

$$0 = \int_a^b \left[F_\lambda(x, u(x), u'(x)) - \frac{d}{dx} F_\xi(x, u(x), u'(x)) \right] dx \\ + F_\xi(b, u(b), u'(b))\varphi(b).$$

Si restringimos nuestra atención a todas las φ que se anulan en b (ya que estas φ tambien son elegibles), concluiriamos como antes, que la ecuación de $E - L$ se debe mantener. Pero una vez tenemos esta información a nuestra disposicion, la identidad anterior implica

$$F_\xi(b, u(b), u'(b))\varphi(b) = 0.$$

Dado que $\varphi(b)$ se puede escoger de manera arbitraria, esto significa que

$$F_\xi(b, u(b), u'(b)) = 0,$$

A esta se le llama la condición de transversalidad o condicion natural en b. Esta es una condición adicional que las soluciones óptimas del problema variacional deben satisfacer. Estas mismas observaciones aplican al extremo izquierdo.

Ejemplo 5.12 *Tratemos de encontrar el valor mínimo que el funcional*

$$I(u) = \frac{1}{2} \int_0^{\log 2} [(u'(x) - 1)^2 + u(x)^2]dx.$$

puede tomar entre todas las funciones u. Esta es una situación en la cual ambos puntos extremos estan libres, asi que como el integrando

$$F(x, \lambda, \xi) = (\xi - 1)^2 + \lambda^2$$

es estrictamente convexo, la solución óptima se encuentra resolviendo el problema

$$u''(x) - u(x) = 0, \quad u'(0) = u'(\log 2) = 1.$$

La solución unica es

$$u(x) = \frac{1}{3}e^x - \frac{2}{3}e^{-x}$$

Otra posibilidad ocurre cuando los valores en los extremos estan restringidos por las desigualdades

$$B_1 \leq u(b) \leq B_2$$

En este caso procedemos de la siguiente manera. Primero examinamos la condición de transversalidad

$$F_\xi(b, u(b), u'(b)) = 0,$$

Si esto determina la solución óptima u de tal manera que es factible, entonces esta es nuestra solución óptima. Si no lo es, es porque $u(b) \notin [B_1, B_2]$, entonces la solución óptima debe exigir que $u(b) = B_1$, o que $u(b) = B_2$, dependiendo si $u(b) < B_1$, o $u(b) > B_2$, respectivamente. Esta regla requiere una explicación mas profunda, pero la tomaremos como valida, y de hecho es correcta en varios casos regulares. Regresaremos a esta discusión mas adelante.

Ejemplo 5.13 *Considere la siguiente situación:*

$$\textit{Minimizar} \quad \int_0^{\log 2} \left[u'(x)^2 + (u(x) - 2)^2\right] dx$$

sujeto a

$$2 \leq u(0) \leq 3, \quad u(\log 2) = 1,$$

La ecuación de E-L junto con las condiciones de frontera nos arroja

$$u''(x) = u(x) - 2, \quad u'(0) = 0, \quad u(\log 2) = 1.$$

Cuya solución es

$$u(x) = 2 - \frac{2}{5}(e^x + e^{-x}).$$

Notamos que $u(0) = 6/5$, asi que esta solución no es admisible para nuestro problema de optimización. Pero, dado que el valor $u(0)$ es menor que el los valores permitidos en 0 concluimos que la solución óptima sera la solución del problema

$$u''(x) = u(x) - 2, \quad u(0) = 2, \quad u(\log 2) = 1$$

La solución óptima es en consecuencia

$$u(x) = -\frac{2}{3}(e^x - e^{-x}) + 2$$

5.5. Problemas variacionales bajo condiciones integrales y puntuales

En esta sección nos gustaria considerar problemas variacionales en los cuales en adicion a tner restricciónes en los valores de los puntos extremos, debemos respetar condiones expresadas en forma de igualdades o desigualdades del tipo

$$\int_a^b G(x, u(x), u'(x))dx \leq \alpha, \quad \int_a^b H(x, u(x), u'(x))dx = \beta$$

Notese que ambos G y H puden ser funciones vectoriales, asi qeu en realidad tenemos varias restricciones integrales. Los vectores α y β estan dados. En esta situación solamente estamos dispuestos a aceptar como funciones factiblea aquellas que respeten todas estas restricciones integrables, y entre ellas nos gustaria encontrar aquella(s) que tenga el menor valor de la integral

$$\int_a^b F(x, u(x), u'(x))dx.$$

Como mencionamos antes tambien podemos tener restricciones en los puntos extremos, note que

$$u(a) = A, \quad u(b) = B$$

es equivalente a

$$u(a) = A, \quad \int_a^b u'(x)dx = B - A.$$

Y asi podemos incorporar la función $G_0(x, \lambda, \xi) = \xi$ como una restricción integral. Entenderemos esta condición de esta manera en lo restante de la sección.

Como se puede esperar de la experiencia que tenemos en programación matematica, tenemos que considerar multiplicadores asociados con todas la restricciones de integraldidad, uno para cada restricción. Asi que tendremos que trabajar con el integrando aumentado

$$\tilde{F}(x, \lambda, \xi) = F(x, \lambda, \xi) + yG(x, \lambda, \xi) + zH(x, \lambda, \xi).$$

El proceso practico de encontrar tales soluciones óptimas es ligeramente diferente de aquellos de programación matematica, aunque esta basado en las mismas ideas fundamentales. Veamos primero en caso de restricciones de igualdad de manera que la función G no esta presente.

Proposición 5.14 *Suponga que hay un vector de numeros z (El vector de multiplicadores) tal que el integrando auxiliar*

$$\tilde{F} = F + zH$$

Resulta ser convexa en (λ, ξ) (es mas, solo se necesita convexidad con respecto a ξ) para cada $x \in (a, b)$ fijo. Si u es la unica solución del problema E-L asociado con tildeF,

$$\frac{d}{dx}\left[\tilde{F}_\xi(x, u(x), u'(x))\right] = \tilde{F}_\lambda(x, u(x), u'(x)),$$

con $u(a) = A$, entonces u es na solución óptima del problema bajo las restricciones integrales correspondientes al vector β determinado por u

$$\beta = \int_a^b H(x, u(x), u'(x))dx.$$

Si la convexidad de $\tilde{F}$ es estricta, entonces se sigue la unicidad de la solución óptima. La prueba de este resultado se reduce a aplicar la parte de convexidad del teorema 5.11a $\tilde{F}$. Los detalles se le dejan al lector.

Este resultado se usa en la practica en dos pasos. Primero, se resuelte la ecuación E-L para $\tilde{F}$ incorporando los multiplicadores (z) durante todo proceso como parametros, de tal manera que obtenemos una familia completa de soluciones, una para cada z . A continuación estos multiplicadores son ajustados d etal manera que la solución óptima nos da el valor apropiado para la restricción integral.

Ejemplo 5.15 *Encuentre la función u que minimiza la integral del cuadrado de su derivada en el intervalo $(0, 1)$ bajo las restricciones*

$$u(0) = u(1) = 0, \quad \int_0^1 u(x)dx = 1.$$

Alternativamente podemos escribir, como mencionamos anteriormente,

$$u(0) = 0, \quad \int_0^1 u'(x)dx = 0, \quad \int_0^1 u(x)dx = 1.$$

El integrando aumentado es ahora

$$\tilde{F}(x, \lambda, \xi) = \xi^2 + z_1\xi + z_2\lambda,$$

con ecuación E-L

$$u''(x) = \frac{z_2}{2}$$

Una primera integración resulta en

$$u'(x) = \frac{z_2}{2}x + c$$

y una segunda integracion teniendo en cuenta que $u(0) = 0$ nos lleva a

$$u(x) = \frac{z_2}{4}x^2 + cx$$

Dado que la función $\tilde{F}$ es siempre estrictamente convexa (con respecto a ξ), la unica solución es encontrada imponiendo la dos restricciones integrales en la anterior función u,

$$0 = \int_0^1 (\frac{z_2}{2}x + c)dx, \quad 1 = \int_0^1 (\frac{z_2}{2}x^2 + cx)dx.$$

Despues de hacer estos calculos obtenemos la solución óptima

$$u(x) = -6x(1 - x).$$

Ejemplo 5.16 *(El cable colgante) Un ejemplo mas interesante de un problema variacioonal bajo restricciones integrales es el siguiente. Nos gustaria determinar la forma adoptada por un cable uniforma colgando de sus puntos extremos a la misma altura bajo la acción de su propio peso, asumiendo que esta forma es el resultado de un proceso de minimización de la energia potencial (Vease Figura 5.6)*

Suponga que colocamos el eje x de tal manera que pase por los dos puntos a la misma altura con una distancia D entre ellos. La longitud del cable es L. Obviamente $L \geq D$. Si ω representa el peso por unidad de longitud,y asumiendo que las secciones transversales son uniformes, la energia potencial asociada co el peso total esta dada por la integral

$$I(u) = w \int_0^D u(x)\sqrt{1 + u'(x)^2}dx,$$

donde como es usual

$$ds = \sqrt{1 + u'(x)^2}dx$$

representa el elemento infinitesimal la longitud de arco. Notese que en este caso u se considera negativo, dado que minimizaremos $I(u)$ entre todas aquellas funciones negativas tal que $u(0) = u(D) = 0$.

Hay otra restricción importante que se debe tener en cuente. Esta es la restricción que dice que la longitud del cable debe ser L. De lo contrario, el problema variacional no tendra ningun sentido fisico, dado que podemos hacer $I(u)$ tan pequeo como deseemos haciendo a u mas negativa. La restriccion que nos referimos es

$$L = \int_0^D \sqrt{1 + u(x)^2}dx.$$

Resumiento, queremos encontrar las soluciones óptimas del problema variacional

$$\textit{Minimizar} \quad \int_0^D u(x)\sqrt{1 + u'(x)}^2$$

sujeto a

$$u(0) = u(D), = 0 \quad L = \int_0^D \sqrt{1 + u'(x)^2}dx.$$

De acuerdo con nuestra discusión anterior, estamos interesados en el integrando

$$F(x, \lambda, \xi) = \lambda\sqrt{1 + \xi^2} + z\sqrt{1 + \xi^2} = (\lambda + z)\sqrt{1 + \xi^2}$$

donde z es el multiplicador asociado a la restricción integral y es visto como un parametro. Dado que formalmente obtenemos el mismo tipo de integrando que en el caso de las superficies mínimas de revolución, no es dificil comprobar que los calculos son formalmente los mismos, asi llegamos a una solución óptima

$$u(x) = c\cosh\left(\frac{x - D/2}{c}\right) - z$$

donde la constante c y el multiplicador z estan determinados por las condiciones

$$z = c\cosh(D/(2c)), \quad L = \int_0^D \sqrt{1+\sinh^2\left(\frac{x-D/2}{c}\right)}dx.$$

La solución es por lo tanto una catenaria.

Ejemplo 5.17 *(El canal) De acuerdo a nuestra discución en el diseño de un canal en el Capítulo 1, el problema consiste en determinar la forma de la sección transversal (una curva) que encierra un area fija tiene mínimo perimetro.*

si u definido en $(0,1)$ *describe un perfil posible le debemos exigir que*

$$u80) = u(1) = 0, \quad A = \int_0^1 u(x)dx,$$

y entre todas aquellas curvas estamos buscando aquella que da el valor mínimo de

$$I(u) = \int_0^1 \sqrt{1+u'(x)^2}dx.$$

Suponga que $u \geq 0$.

Como antes debemos trabajar con la ecuación E-L para la función

$$F(x,\lambda,\xi) = \sqrt{1+\xi^2} + z\lambda,$$

donde z es el multiplicador. Notese que esta función es convexa en (λ,ξ)*, afín en* λ*, y estrictamente convexa en* ξ*. Esto es suficiente para garantizar la unicidad de la solución óptima (revisese la prueba del terema 5.11)*

$$\left(\frac{u'(x)}{\sqrt{1+u'(x)^2}}\right) = z.$$

Despues de un par de cálculos elementales tenemos

$$u'(x) = \frac{zx+c}{\sqrt{1-(zx+c)^2}},$$

donde c es una constante. Entonces

$$\begin{aligned} u(x) &= \int_0^x \frac{zs+c}{\sqrt{1-(zs+c)^2}} \\ &= -\tfrac{1}{z}\sqrt{1-(zs+c)^2}|_0^x \\ &= \tfrac{1}{z}\left(\sqrt{1-c^2} - \sqrt{1-(zx+c^2)}\right) \end{aligned}$$

La condición $u(1) = 0$ *conduce a* $c = -z/2$ *y por lo tanto*

$$\frac{1}{2z}\left(\sqrt{4-z^2} - \sqrt{4-z^2(2x-1)^2}\right)$$

El multiplicador z se determina de manera que satisfaga la restricción integral pero esto lleva a unos cálculos complicados. Lo que es importante es notar que esta solución óptima u es el arco de un círculo.

Ahora trataremos el caso de las restricciones integrales en la forma de igualdades y desigualdades de manera simultanea. Nos enfocaremos en el problema

$$\text{Minimizar} \quad I(u) = \int_a^b F(x, u(x), u'(x))dx$$

sujeto a

$$\begin{aligned} u(a) &= a \\ \int_a^b G(x, u(x), u'(x)dx &\leq \alpha, \\ \int_a^b H(x, u(x), u'(x)dx &= \beta. \end{aligned}$$

De nuevo nuestra experiencia con programación no lineal hace el siguiente resultado plausible.

Proposición 5.18 *Suponga que hay un vector (y, z), $y \geq 0$, de tal manera que la función*

$$\tilde{F}(x, \lambda, \xi) = F(x, \lambda, \xi) + yG(x, \lambda, \xi) + zH(x, \lambda, \xi)$$

es convexa en (λ, ξ) (de nuevo la convexidad con respecto a ξ es suficiente). Si v es la solución (unica) del problema E-L correspondiente,

$$\frac{d}{dx}\left[\tilde{F}_\xi(x, v(x), v'(x))\right] = \tilde{F}_\lambda(x, v(x), v'(x)), v(a) = A,$$

entonces v es la solución óptima del problema variacional anterior, dado que

$$\begin{aligned} \int_a^b G(x, v(x), v'(x)) &\leq \alpha, \\ \int_a^b H(x, v(x), v'(x))dx &= \beta \\ y\left(\int_a^b G(x, v(x), v'(x))dx - \alpha\right) &= 0. \end{aligned}$$

Resolvamos un ejemplo

Ejemplo 5.19 *Resolver*

$$\textit{Minimizar} \quad \int_0^1 u'(x)^2 dx$$

suejeto a

$$u(0) = u(1) = 1, \int u(x)^2 dx \leq \alpha,$$

donde α es un numero no negativo dado. Introduciendo un multiplicador y para encarganos de la restricción inetegral, debemos enfrentarnos a la ecuación $E-L$ para el integrando aumentado

$$\tilde{F}(x,\lambda,\xi)=\xi^2+y\lambda^2$$

viendo a y como un parametro. Este problema consiste en

$$u''(x)=yu(x),\quad u(0)=0,u(1)=1.$$

La solución general es de la forma

$$u(x)=\frac{\sinh(\sqrt{y}x)}{\sqrt{y}}$$

si y no se anula, y

$$u(x)=x$$

en caso que $y=0$. Cada vez que α sea escogido de tal manera que

$$\int_0^1 x^2dx=\frac{1}{3}\leq\alpha,$$

la solución óptima sera la línea $u8x)=x$. Pero si

$$\alpha<\frac{1}{3}$$

la solución óptima sera de la forma

$$u(x)=\frac{\sinh(\sqrt{y}x)}{\sinh(\sqrt{y})},$$

donde el multiplicador y esta determinado de tal manera que

$$\int_0^1\left(\frac{\sinh(\sqrt{y}x)}{\sinh(\sqrt{y})}\right)^2=\alpha.$$

Cuando se imponen restricciones adicionales para problemas variaciones en forma puntual como

$$H(x,u(x),u'(x))=0$$

o

$$G(x,u(x),u'(x))\leq 0,\quad H(x,u(x),u'(x))=0$$

entonces los multiplicadores $y(x),z(x)$ deben ser funciones de x, porque debemos satisfacer una restricción para cada x. De esta manera el funcional que nos deja tratar este tipo de restriccion puntual es

$$I(u,y,z)=\int_a^b[F(x,u(x),u'(x))+y(x)G(x,u(x),u'(x))+z(x)H(x,u(x),u'(x))]dx.$$

Note la dependencia explicita de I con respecto a los multiplicadores $y(x),z(x)$.

No hay duda qe una de las situaciones mas importantes en las cuales las restricciones puntuales se tienen que tener en cuenta es aquellas de problemas de control óptimo. Dado que el ultimo capítulo de este texto esta dedicado a estos y dada su importancia no desarrollaremos el tema en este momento.

5.6. Resumen de las restricciones para problemas variacionales

El objetivo de esta sección es clarificar todas las posibilidades que pueden surgir cuando se considera un típico problema variacional desde la perspectiva de distintos tipos de restricciones. Esta discución considera, como casos particulares, condiciones de transversalidad de todos los tipos, y todas las situaciaciones ralacionadas con restricciones integrales y restricciones en los extremos.

Considere el problema

$$\text{Minimizar} \quad \int_a^b F(x, u(x), u'(x))dx$$

sujeto a

$$\int_a^b G(x, u(x), u'(x))dx \leq \alpha \quad \int_a^b H(x, u(x), u'(x))dx = \beta, \quad u(c) \leq A,$$

donde $c = a$ o $c = b$ esta fijo.

Sabemos que para encontrar una solución óptima debemos resolver E-L para el integrando aumentado

$$\tilde{F} = F + yG + zH$$

Suponga que $\tilde{u}(x, y, z, w)$ es la solución general de la ecuación E-L para $\tilde{F}$ donde $w = (w_1, w_2)$ representan dos condiciones arbitrarias de integración , desde que la ecuación E-L es de segundo orden. Defina

$$\begin{aligned} f(y, z, w) &= \int_a^b F(x, \tilde{u}(x, y, z, w), \tilde{u}'(x, y, z, w))dx, \\ g(y, z, w) &= \int_a^b G(x, \tilde{u}(x, y, z, w), \tilde{u}'(x, y, z, w))dx, \\ h(y, z, w) &= \int_a^b H(x, \tilde{u}(x, y, z, w), \tilde{u}'(x, y, z, w))dx, \\ \varphi(y, z, w) &= \tilde{u}(c, y, z, w) \end{aligned}$$

Entonces no es difícil convercernos que los valores óptimos para (y, z, w) deben ser determinados resolviendo el PPNL

$$\text{Minimizar} \quad f(y, z, w)$$

sujeto a

$$\begin{aligned} g(y, z, w) \leq \alpha, \quad y \geq 0, \quad y(g(y, z, w) - \alpha) = 0, \\ h(y, z, w) = \beta, \quad \varphi(y, z, w) \leq A. \end{aligned}$$

Una vez se han encontrado estos valores óptimos (y_0, z_0, w_0) es importante regresar a $\tilde{F}$ "a posteriori" y confirmar que

$$\tilde{F} = F + y_0 G + z_0 H$$

es convexa con respecto a ξ. En particular, dado que $y \geq 0$, la función G (o todos sus componentes) debe ser convexa con respecto a ξ. si $\tilde{F}$ es no convexo, entonces podemos no tener la solución óptima. En este contexto, tambien podemos tratar todas las variables relacionadas a diferentes tipos de igualdades y/o desigualdades.

Esta perspectiva general lleva en algunas ocasiones a problemas que no son suaves o incluso en algunos casos que no son continuos y, que requeririan en principio, tecnicas mas elaboradas para encontrar soluciones óptimas. Dada la estructura especial de las restricciones, es elemental notar, que como mencionamos anteriormente cuando trabajamos con las condiciones de optimalidad de un PPNL, las restricciones

$$g(y, z, w) \leq \alpha, \quad y \geq 0, \quad y(g(y, z, w) - \alpha) = 0$$

nos lleva a tener

$$y_i(g_i(y, z, w) - \alpha_i) = 0$$

para todos los i, y estas ecuaciones nos permiten un tratamiento separado de varios PPNL. Por ejemplo, cuando G tiene una sola componente entonces, tendriamos dos posibilides,

$$y = 0 \quad \text{o} \quad g(y, z, w) = \alpha,$$

que llevan a los dos PPNL

$$\text{Minimizar} \quad f(o, z, w)$$

sujeto a

$$g(0, z, w) \leq \alpha, \quad h(o, z, w) = \beta, \quad \varphi(0, z, w) \leq A$$

y

$$\text{Minimizar} \quad f(y, z, w)$$

sujeto a

$$\text{Minimizar} \quad f(y, z, w)$$

sujeto a

$$g(y, z, w) = \alpha, \quad h(y, z, w) = \beta, \quad \varphi(y, z, w) \leq A,$$

consideranto estas soluciones óptimas para $y \geq 0$. La verdadera solución óptima de nuestro problema se encontrara en alguno de estos dos PPNL.

Ejemplo 5.20 *Nos gustaria resolver*

$$\textit{Minimizar} \quad \frac{1}{2}\int_0^1 (u'(x) - 1)^2 dx$$

suejto a

$$o \leq u(0) \leq, \quad 0 \leq u(1) \leq \frac{1}{2}$$

La ecuación E-L es $u'' = 0$, asi que la solución general es

$$u(x) = w_1 x + w_2.$$

Por lo tanto, consideramos la función objetivo

$$f(w) = \frac{1}{2}\int_0^1 (w_1 - 1)^2 dx = \frac{1}{2}(w_1 - 1)^2.$$

Las restricciones corresponden a requerir

$$0 \leq u(0) = w_2 \leq 1, \quad 0 \leq u(1) = w_1 + w_2 \leq \frac{1}{2}$$

Por lo que debemos resolver PPNL

$$\textit{Minimizar} \quad \frac{1}{2}(w_1 - 1)^2$$

bajo las restricciones

$$0 \leq w_1 \leq 1, \quad 0 \leq w_1 + w_2 \leq \frac{1}{2}$$

Es muy facil encontrar (incluso graficamente) la solución óptima, que corresponde a $(1/2, 0)$, asi que la solución óptima a nuestro problema es $u(x) = x/2$

Ejemplo 5.21 *Considere el problema*

$$\textit{Minimizar} \quad \frac{1}{2}\int_0^1 u'(x)^2 dx$$

bajo las restricciones

$$0 \leq u(0) \leq 1, \quad u(1) = 1, \quad \int_0^1 u(x)dx \leq \frac{1}{2}$$

La ecuación E-L que debemos resolver es $u'' = y$, cuya solución general es

$$u(x) = y\frac{x^2}{2} + w_1 x + w_2.$$

Es muy facil obtener

$$f(y, w) = \frac{1}{6}(y^2 + 3yw_1 + 3w_1^2),$$
$$g(y, w) = \frac{y}{6} + \frac{w_1}{2} + w_2, \varphi(y, w) = \left(w_2, w_1 + w_2 + \frac{y}{2}\right).$$

Entonce sel PPNL a ser considerado es

$$\textit{Minimizar} \quad \frac{1}{6}(y^2 + 3yw_1 + 3w_1^2)$$

Sujeto a

$$\frac{y}{6}+\frac{w_1}{2}+w_2\leq\frac{1}{2},\quad w_1+w_2+\frac{y}{2}=1,\quad 0\leq w_2\leq 1,$$
$$y\geq 0,\quad y\left(\frac{y}{6}+\frac{w_1}{2}+w_2-\frac{1}{2}\right)=0.$$

Como se menciono anteriormente este PPNL se divide en

$$\textit{Minimizar}\quad \frac{w_1^2}{2}$$

sujeto a

$$\frac{w_1}{2}+w_2\leq\frac{1}{2},\quad w_1+w_2=1,\quad 0\leq w_2\leq 1,$$

Con $(0,1)$ *como unico punto admisible con costo asociado* $1/2$*, y*

$$\textit{Minimizar}\quad \frac{1}{6}(y^2+3yw_1+3w_1^2)$$

sujeto a

$$\frac{y}{6}+\frac{w_1}{2}+w_2=\frac{1}{2},\quad w_1+w_2+\frac{y}{2}=1,\quad 0\leq w_2\leq 1,$$

Con $y\geq 0$*. Usando las dos restricciones lineales obtenemos* $w_1=1-2y/3$*,* $w_2=y/6$*, u la condición* $0\leq w_2\leq 1$ *lleva a* $0\leq y\leq 6$*. Colocando esto junto y haciendo las sustituciones apropiadas, estamos interesados en encontrar el mínimo de la parabola*

$$\frac{1}{6}\left(\frac{y^2}{3}-y+3\right)$$

en el intervalo $0\leq y\leq 6$*. Este mínimo se alcanza en* $y=3/2$ *con costo* $3/8$ *y* $w_1=0$*,* $w_2=1/4$*. Desde que este costo óptimo es mas pequeño que el encontrado en el subproblema anterior, concluimos que la solución óptima buscada es*

$$u(x)=\frac{3x^2+1}{4}$$

5.7. Problemas variacionales de diferente orden

En algunos casos pdemos estar interesados en problemas variacionales donde aparece la segunda derivada (o derivadas de orden superior) o donde ninguna derivada esta presente. El orden mas alto de las derivadas presentes en un problema variacional es el orden del problema. Hasta ahora, nos hemos concentrado en problemas de primer orden. Estos son los mas comunes por un amplio margen. Pero nos gustaria mencionar algunas cosas acerca de los problemas variacionales de orden cero y dos.

Los problemas variacionales de orden cero, i.e, aquellos en los que no aparecen derivadas en el funcional de costo, estan incluidos en nuestra discución

previa, porque al fin y al cabo estos son un caso especial de los problemas de primer orden donde no hay dependencia en las primeras derivadas. En estos casos la ecuación E-L no es una ecuación diferencial, sino una ecuación algebraica. Resolver esta ecuación nos dará la solución óptima despues de ajustar las constantes con los valores en los puntos extremos. Un ejemplo aclarara lo que decimos.

Ejemplo 5.22 *Un contenedor cilindrico rota alrededor de su eje a una velocidad constante ω_0 (Figura 5.7). Nos gustaria determinar el perfil adoptado por cierto fluido en su interior, asumiendo que la forma es el resultado de un proceso de minimización de la energía potencial.*

Especificamente, y debido a la simetria radial, si

$$z(r), \quad 0 \leq r \leq R$$

describe un perfil dado, la energia potencial asociado con el esta dada por la integral

$$U(z) = \int_0^R \pi\rho[gr(z(r)^2 - 2Hz(r)) - \omega_0^2 r^3 z(r)]dr,$$

donde g, ρ y H son constantes. Obviamente, el perfil $z(r)$ debe respetar la restricción de volumen

$$V = 2\pi \int_0^H rz(r)dr$$

para una constante fija V. En resumen buscamos

$$\textit{Minimizar} \quad \int_0^H [gr(z(r)^2 - 2Hz(r)) - w_0 r^3 z(r)]dr$$

sujeto a

$$V = 2\pi \int_0^H rz(r)dr.$$

Note que la derivada $z'(r)$ no aparece en el funcional de costo.

Si introduciomos un multiplicador λ para tener en cuenta la restricción de volumen, tenemos que escribir la ecuación E-L para la función

$$F(x, u) = [g(u^2 - 2Hu)r - \omega - 0^2 r^3 u] + \lambda ru.$$

Esto es

$$0 = \frac{\partial F}{\partial u}(r, z(r)),$$

i.e,

$$g(2z(r) - 2H)r - \omega_0^2 r^3 + \lambda r = 0.$$

Resolviendo para $z(r)$, obtenemos

$$z(r) = \frac{\omega_0^2}{2g} r^2 + H - \frac{\lambda}{2g}.$$

A través de la restricción de volumen determinamos la constante

$$H - \frac{\lambda}{2g}$$

De hecho tenemos

$$H - \frac{\lambda}{2g} = \frac{V}{\pi R^2} - \frac{w_0^2}{4g} R^2,$$

y el perfil óptimo es

$$z(r) = \frac{\omega_0^2}{2g}\left(r^2 - \frac{R^2}{2}\right) + \frac{V}{\pi R^2}$$

note que esto es una parabola rotada alrededor del eje del cilindro.

Los problemas de segundo orden son mas elaborados, como uno podria anticipar. La estrategia para encontrar condiciondes de optimalidad en la forma de ecuaciondes E-L es sin embargo, similar a la de los problemas de primer orden. El ecuación E-L sera una ecuación diferencial de cuarto orden, la cual sera complementada con condiciones de fontera apropiadas. Le dejamos al lector (aunque incluimos en la discución mas adelante) justificar el siguiente hecho.

Teorema 5.23 *Si el integrando para el problema de segundo orden es*

$$F(x, u, u', u'').$$

entonces la ecuación E-L es

$$\frac{d^2}{dx^2}\frac{\partial F}{\partial u''} - \frac{d}{dx}\frac{\partial F}{\partial u'} + \frac{\partial F}{\partial u} = 0$$

Las condiciones en los puntos extremos pueden incluir valores de u y/o u'. Tambien tenemos que tener en cuenta la transversalidad o condiciones naturales de frontera.

Muchos problemas relacionados con el doblamiento o torcimiento de barras elasticas delgadas estan formulados como problemas variacionales de segundo orden, la causa de esto es que las energias dependen en las curvaturas (segundas derivadas) de los perfiles adoptados. A continuación un ejemplo elemental.

Ejemplo 5.24 *Una barra delgada y elastica es doblada como se muestra en la figura 5.8. Si*

$$u(x), \quad x \in (0, L).$$

describe la linea central de la barra, la energia potencial acumulada en tal estado esta dada por la integral

$$P(u) = \int_0^L \frac{u''(x)^2}{(1 + u'(x)^2)^{5/2}} dx$$

donde $k > 0$ es una constante conocida. Las condiciones en los puntos extremos son

$$u(0 = u'(0) = 0, \quad u(L) = L_1$$

Dado que hemos impuesto tres condiciones en los puntos extremos, necesitamos otra dado que la ecuación E-L tendra orden 4. Esta condicion es una condición natural de frontera en el punto derecho L. Dado que la condicion de frontera que falta es para $u'(L)$, la condición de transversalidad que necesitamos es el factor con $u'(L)$ cuando integramos por partes analizando la función

$$g(t) = \int_0^L F(x, u(x) + t\varphi(x), u'(x) + t\varphi'(x), u''(x) + t\varphi''(x))dx,$$

y exigiendo que

$$0 = g'(x) = \int_0^L \left(\frac{\partial F}{\partial u}\varphi + \frac{\partial F}{\partial u'}\varphi' + \frac{\partial F}{\partial u''}\varphi'' \right) dx$$

Integrando dos veces por partes y teniendo en cuenta las otras condiciones en los puntos extremos obtenemos

$$0 = \int_0^L \varphi \left(\frac{d^2}{dx^2}\frac{\partial F}{\partial u''} - \frac{d}{dx}\frac{\partial F}{\partial u'} + \frac{\partial F}{\partial u} \right) dx + \varphi'(L)\frac{\partial F}{\partial u''}(L, u(L), u'(L), u''(L)).$$

Esto nos da información acerca de la ecuación E-L, y ademas nos dice que la condición natural que estamos buscando es

$$\frac{\partial F}{\partial u''}(L, L_1, u'(L), u''(L)) = 0.$$

En nuestro ejemplo particular esta condición es

$$u''(L) = 0$$

Despues de escribir cuidadosamente E-L, debemos resolver el problema

$$\frac{d^2}{dx^2}\frac{2u''}{(1+u'^2)^{5/2}} + \frac{d}{dx}\frac{5(u'')^2u'}{(1+u'^2)^{7/2}} = 0,$$
$$u(0) = u'(0) = 0, \quad u(L) = L_1, \quad u''(L) = 0.$$

Esta ecuación es imposible de resolver a mano. Una aproximación razonable sera indicar que el perfil esperado tendra una derivada pequeña u', asi que podemos ignorar los terminos que incluyan u'. Esta simplificación, junto con una integración nos llevan a

$$u''' + \frac{5}{2}(u'')^2u' = constante,$$

y es mas

$$u''' = constante$$

Esto junto con las cuatro condiciones de frontera implica que

$$u(x) = -\frac{L_1}{2L^3}x^3 + \frac{3L_1}{2L^2}x^2$$

nos da una aproximación razonable del perfil que toma la barra

5.8. Programación Dinamica: La ecuación de Bellman

En varia situaciones practica de interes, se desea mover un sistema sucesivamente a traves de un numero de diferentes pasos para completar un proceso completo y llegar a un estado deseado. Cada una de estas acciones tiene un costo asociado. Dado un objetivo especifico que se desea alcanzar dado un estado inicial, nos gustaria determinar la estategia global óptima que tiene el menor costo asociado.

Sea t una variable que indica las etapas sucesivas en las cuales se debe tomar una desición acerca de como dirigir el sistema

$$t = t_i, \quad i = 0, 1, \ldots, n;$$

sea x la variable que describe el estado del sistema. En cada paso i, nos gustaria tener

$$x \in A_i$$

si A_i es el conjunto (finito) de estados factibles cuanto $t = t_i$. El costo asociado con pasar de $x \in A_i$ a $y \in A_{i+1}$ esta dado por

$$c(i, x, y)$$

dado un estado inicial (t_0, x_0), estamos interesados en la tarea de determinar la estrategia óptima para alcanzar el esto final deseado (t_n, x_n) con el menor costo. Esta es una situación típica de programación dinamica.

Asuma que para $0 < j < n$ conocemos el camino óptimo que comienza en (t_0, x_0) y va por (t_j, x) para cada $x \in A_j$. Si

$$S(t_j, x)$$

es el costo asociado con tal estrategia óptima terminando en (t_j, x). ¿Cómo podemos encontrar la solución óptima comenzando en (t_0, x_0) y terminano en (t_{j+1}, y) para cualquier $y \in A_{j+1}$ dada? No es dificil convercernos que debemos resolver el problema

$$\min_{x \in A_j} [S(t_j, x) + c(j, x, y)]$$

Esta es el principio fundamental o propiedad de la programación dinamica, y a traves de el podemos encontrar la estrategia óptima desde (t_0, x_0) a (t_n, x_n) de la manera mas racional

Proposición 5.25 *(Propiedad fundamental de la programación dinamica) Si $S(t_j, x)$ denota el costo óptimo desde (t_0, x_0) a (t, j, x), entonces debemos tener*

$$S(t_{j+1}, y) = \min_{x \in A_j} [S(t_j, x) + c(j, x, y)].$$

A continuación una situación simplificada.

Ejemplo 5.26 *(El viajero) Un pasajero quiere ir desde la ciudad A a la ciudad H por el camino mas corto de acuerdo al mapa en la figura 5.9, donde los numeros indican distancias entre las correspondientes ciudades.*

Evidentemente podemos hacer un conteo exhaustivo de todas las posibilidades y decidir cúal es la mejor. En esta situación simplificada esto puede no ser una mala idea. Sin embargo, nos gustaria ilustrar la propiedad fundamental de la programación dinamica en este ejemplo. De acuerdo a la proposición 5.25, debemos proceder sucesivamente determinando $S(t_j, x)$ para cada $x \in A_j$ para terminar con $S(t_n, x_n)$. En el ejemplo propuesto, tenemos cuatro estados t_0, t_1, t_2, t_3, t_4 con conjuntos asociados de estados factibles

$$A_0 = \{A\}, \quad A_1 = \{B, C, D\}, \quad A_2 = \{E, F, G\}, \quad A_3 = \{H\}.$$

Para cada ciudad en A_1 hay un unico camino desde A, asi que este debe ser óptimo y

$$S(t_1, B) = 7, \quad S(t_1, C) = 4, \quad S(t_1, D) = 1.$$

Para cada ciudad en A_2 determinamos el costo óptimo basados en la propiedad fundamental de programación dinamica,

$$S(t_{j+1}, y) = \min_{x \in A_j} [S(t_j, x) + c(j, x, y)].$$

Es nuestro ejemplo concreto estamos buscando el mínimo de

$$\begin{array}{lll} 7+4, & 4+4, & 1+8, \\ 7+6, & 4+5, & 1+4, \\ 7+2, & 4+7, & 1+5, \end{array}$$

El ultimo paso nos lleva al camino mas corto deseado. Nos gustaria encontrar el mínimo de

$$\min\{8+3, 5+6, 6+2\} = 8.$$

Por lo tanto, la distancia mas corta es 8 y corresponde a la ruta $A-D-G-H$. Notese que cuando se usa la propiedad fundamental de programación dinamica siempre tenemos que operar con las sumas de dps números, mientras que un conteo directo exhaustivo requeriria (en esta situación simplificada) trabajar con las sumas de tres números. No es dificil inferir la importancia de este hecho para problemas mas complicados,

Ejemplo 5.27 *Otro ejemplo típico en programacion dinamica es aquel de una compañia que puede producir tres tipos de productos de la leche: queso, mantequilla y yogur. El beneficio de estos productos usando $1, 2, 3,$ o 4 unidades de leche esta dada en la tabla 5.10.*

¿Cúal es el máximo beneficio que puede ser obtenido con 4 unidades de leche? De nuevo, en esta situación simplificada, no es dificil encontrar la solución mediante un analisis exhaustivo de todas las posiblidades. El patron en el contexto

de programación dinamica es como sigue: Identificamos cada uno de los productos con los valores de la variable t, t_0, t_1 *y* t_2. *En conjunto de estados posibles en cada paso será* $\{0, 1, 2, 3, 4\}$, *lo que quiere decur qye podemos asignar cada uno de estos números de unidades de leche a cada uno de los tres productos, mientras no exedamos las* 4 *unidades posibles. Usando los datos en la tabla 5.10, obtenemos los resultados de la tabla 5.11.*

El máximo beneficio es 28, *y corresponde a* 3 *unidades de leche para queso y una unidad para yogur.*

El principio basico de programación dinamica aplicado al caso continuo mps da otra perspectiva en los problemas variacionales en la cual nos enfocamos en los valores óptimos en ves de las soluciones óptimas. Definimos la "función valor"'poniendo

$$S(t,x) = \min_u \left\{ \int_t^T F(\tau, u(\tau), u'(\tau))d\tau : u(t) = x, u(T) = B \right\},$$

donde T y B son datos fijos. Entonces $S(t,x)$ nos da el costo óptimo asociado con el problema comenzando en (t,x) y terminando en (T,B). En la aproximación variacional inistimos en caminos óptimos o soluciones óptimas, y no tanto en los valores óptimos

La propiedad basica de esta función valor $S(t,x)$ es precisamente el principio fundamental de programación dinamica, ya mencionado para el caso discreto, el cual en ete contexto se puede escribir como

$$\min_x \left\{ \min_u \left\{ \int_t^{t'} F(\tau, u(\tau), u'(\tau))d\tau : u(t) = x, u(t') = z \right\} + S(t', z) \right\},$$

donde $t < t' < T$ (vease figura 5.12)

Esta condición puede ser reorganizada como sigue:

$$0 = \min_z \left\{ \min_u \left\{ \frac{1}{t'-t} \int_t^{t'} F(\tau, u(\tau), u'(\tau))d\tau : u(t) = x, u(t') = z \right\} + \frac{S(t', z) - S(t, x)}{t' - t} \right\}.$$

Si hacemos que $z = x + t(t - t')$, el mínimo de z puede ser transformado en un mínimo en y, y

$$0 = \min_y \left\{ \min_u \left\{ \frac{1}{t-t'} \int_t^{t'} F(\tau, u(\tau), u'(\tau))d\tau : u(t) = x, u(t') = x + y(t' - t) \right\} + \frac{S(t', x + y(t' - t)) - S(t, x)}{t - t'} \right\}.$$

¿Qué pase si dejamos que $t' \searrow t$? Para cada función u tal que $u(t) = x$ y $u(t') = x + y(t' - t)$, por el teorema fundamental del cálculo tenemos

$$\frac{1}{t'-t} \int_t^{t'} F(\tau, u(\tau), u'(\tau))d\tau \to F(t, x, y),$$

mientras que si asumimos que que S es diferenciable, entonces por la regla de la cadena,

$$\frac{S(t', x + y(t' - t)) - S(t, x)}{t' - t} \to \frac{\partial S}{\partial t}(t, x) + y\frac{\partial S}{\partial x}(t, x).$$

Concluimos que

$$0 = \min_y \left[F(t, x, y) + \frac{\partial S}{\partial t}(t, x) + y\frac{\partial S}{\partial x}(t, x) \right],$$

o

$$-\frac{\partial S}{\partial t}(t, x) = \min_y \left[F(t, x, y) + y\frac{\partial S}{\partial x}(t, x) \right]$$

.

Esta es la ecuación de Bellman de programación dinamica. Hemos incluido esta derivación informal de esta porque seguiremos un camino similar para establecer el principio del máximo de Pontryagin para problemas de control óptimo en el siguiente capítulo.

Esta derivación puede llevar, sin mucha dificultad, a las ecuaciones E-L para soluciónes óptimas de problemas variacionales. Pero como ya mencionamos, esta sera nuestra estrategia principal para condiciones necesaria para la optimalidad no isistimos en esto aquí.

Ejemplo 5.28 *En algunas simplificaciones simples, los calculos para obtener la ecuación de Bellman se pueden llevar a cabo de manera explicita. Suponga que $F(t, \lambda, \xi) = \xi^2$. Dado que esta función es estrictamente convexa y depende solamente de la derivada, sabemos que la solución óptima u que satisface $u(t) = x, u(T) = B$ es la función lineal con pendiente $(B - x)/(T - t)$. Por lo tanto el valor de la función es*

$$S(t, x) = (T - t)\left(\frac{B - x}{T - t}\right)^2 = \frac{(B - x)^2}{T - t}.$$

Por otro lado para toda constante α,

$$\min_y (y^2 + y\alpha) = -\frac{\alpha^2}{4},$$

asi que la ecuación de Bellman es

$$4\frac{\partial S}{\partial t} = \left(\frac{\partial S}{\partial x}\right)^2.$$

Es un ejercicio sencillo confirmas que la forma explicita de $S(t, x)$ satisface esta ecuación diferencial parcial.

5.9. Ideas basicas en la aproximación númerica

Es casi evidente que las soluciones óptimas para muchos problemas no pueden ser resueltas analíticamente. Las tecnicas de aproximación son por lo tanto una herramienta indispensable para resolver muchos problemas variacionales. La idea basica de la aproximación numerica de los problemas continuos de optimización es "discretización".

Dado un problema variacional, debemos construir una versión discretizada del mismo con cierto nivel de presición que esta relacionado con la finura de la discretación que hemos utilizado. Tal versión discretizada sera ahora un problema de programación al cual le podemos aplicar todos los algoritmos computacionales descritos en el capítulo 4. Desde este punto de visto los algoritmos númericos son el puente entre los problemas de optimización de dimensión finita e infinita.

Otra posibilidad interesante para aproximar soluciones óptmas a problemas variacionales es explotar las ecuaciones E-L. Pero ya que este es un libro de optimización, y un este acercamiento nos llevara a aproximar ecuaciones diferenciales, nos apegaremos a los conceptos y tecnicas genuinos de optimización. Es por lo tanto importante siempre tener en mente cualquier información valiosa acerca del problema cuya solución tratamos de aproximar.

Suponga que tenemos un típico problema variacional

$$\text{Minimizar} \quad I(u) = \int_a^b F(x, u(x), u'(x))dx$$

con $u(a) = A$, $u(b) = B$. Usualmente, las discretizaciones de las integrales se hacen dividiendo el intervalo $[a, b]$ en un cierto numero de subintervalos, $n + 1$, y suponga por ahora que los funciones factibles para el nuevo problema discretizado son afines a trozos, i.e, son afines en cada subintervalo

$$\left[a + j\frac{(b-a)}{(n+1)}, a + (j+1)\frac{(b-a)}{(n+1)}\right] \tag{5.1}$$

Note que tales funciones estan unicamente determinadas por sus valores en los nodos

$$a + j\frac{(b-a)}{(n+1)}, \quad j = 1, \ldots, n$$

y por lo tanto los vectores factibles para este nuevo problema de optimización corresponderan a estos valores. Vemos que este proceso cambia el problema original infinito dimensional, a un problema en dimensión finita. La idea es que haciendo el numero de subintervalos (n+1) mas grande, las soluciones óptimas para estos problemas discretizados se pareceran cada vez mas y aproximaran con mayor presición bajo condiciones que ignoraremos aquí, a la verdadera solución óptima. Sean

$$X = (x_j)_{1 \leq j \leq n} \tag{5.2}$$

los valores nodales de las funciones factibles. De esta manera la función u que consideraremos para $I(u)$ será

$$u(x) = A + \sum_{k=1}^{j} \frac{(x_{k+1} - x_k}{n+1} + (x_{j+1} - x_j) \left(x - a - k\frac{b-a}{n+1} \right)$$

si

$$x \in \left[a + k\frac{b-a}{n+1}, a + (k+1)\frac{b-a}{n+1} \right]$$

Esta es la función continua afin a trozos que toma los valores x_j en los puntos nodales $a+j(b-a)/(n+1)$. Hay una forma útil de expresar esta función como una combinación lineal de ciertas "funciones basicas". Expresamente, si definimos

$$\psi_{j,n}(x), \quad j = 0, 1, \ldots, n, n+1$$

cono la función afin a trozos cuyo valor en los nodos x_i para $i \neq j$ es cero, y precisamente en x_j es la unidad (vease figura 5.13), entonce esclaro que la funcion u con valores nodales x_j se pude escribir como

$$u(x) = \sum_j x_j \psi_{j,n}(x) \tag{5.3}$$

Para la derivada tenemos

$$u'(x) = \sum_j x_j \psi'_{j,n}(x).$$

De esta forma es inmediato como escribir un problema de optimización 'para el vector X en 5.2 calculando $I(u)$ para u como en 5.3. Esta sera la versión discreta y aproximada para nuestro problema variacional. Especificamente

$$T(X) = I(u) = \int_a^b F(x, u(x), u'(x))dx$$
$$= \int_a^b F\left(x, \sum_j x_j \psi_{j,n}(x), \sum_j x_j \psi'_{j,n}(x) \right) dx.$$

Si partimos esta integral en una suma sobre los subintervalos 5.1, y notamos que cada $\psi_{j,n}$ es lineal, de tal manera que su derivada es constante en ellos, reuniendo todas las contribuciones de un intervalo en particular podemos escribir de manera mas explicita

$$T(X) = \sum_{j=0}^{n} \int_{a+j(b-a)/(n+1)}^{a+(j+1)(b-a)/(n+1)} F\left(x, \sum_j x_j \psi_{j,n}(x), (n+1)(x_{j+1} - x_j) \right) dx.$$

Es mas, podemos usar una cuadratura simple para aproximar estas integrales. La forma final usando la regla trapezoidal es

$$T(X) = \sum_{j=0}^{n} \frac{1}{2(n+1)} \left[F\left(a + (j+1)\frac{b-a}{n+1}, x_{j+1}, (n+1)(x_{j+1} - x_j)\right) \right.$$
$$\left. + F\left(a + j\frac{b-a}{n+1}, x_j, (n+1)(x_{j+1} - x_j)\right) \right],$$

donde $x_0 = A$, $x_{n+1} = B$. En terminos del vector X de valores nodales, nos enfrentamos con problema de programación (no lineal, sin restricciones). Resolviendolo, obtenemos una solución aproximada al problema inicial continuo de optimización. La forma del funcional $T(x) = I(u)$ in terminos de X depende de cada situación particular

Ejemplo 5.29 *Para la superficie mínima de revolución tomamos*

$$I(u) = \int_0^1 u(x)\sqrt{1 + u'(x)^2}dx,$$
$$u(0) = 1, \quad u(1) = 1.$$

La función objetivo resultante $T(X)$ en terminos de los valores nodales como se indico en la discución anterior es

$$T(X) = \sum_{j=0}^{n} \sqrt{1 + ((n+1)(x_{j+1} - x_j))^2}\frac{x_{j+1} + x_j}{2(n+1)}.$$

Teniendo en cuenta que $x_0 = 1$, $x_{n+1} = 1$, vemos que este es el funcional que tenemos que minimizar con ayuda de los algoritmos del capítulo 4, para varios valores del número de subintervalos n. Haciendo esto, encontramos que estos concuerdan bastante bien con el arco de una catenaria, que es la solución óptima del problema continuo de optimización.(Figura 5.14)

De manera alternativa podemos considerar la discretización considerando como variables independientes las pendientes en cada subintervalo. Esto conduce a una forma mas simple del funciónal objetivo, pero tendriamos que forzar una restricción (lineal), dado que las pendientes en los diferentes subintervalos deben ser ral que el valor de u en 1 esta dado, y esto impone una restricción en el conjunto de posibles restricciones. En vez de resolver de nuevo el mismo ejemplo en este formato con una restricción integral preferimos comenzar otro ejemplo.

Ejemplo 5.30 *El problema para el diseño del canal es*

$$\textit{Minimizar} \quad \int_0^1 \sqrt{1 + u'(x)^2}dx$$

sujeto a

$$u(0) = u(1) = 0, \quad \frac{1}{3} = \int_0^1 u(x)dx.$$

Como se hizo con el ejemplo anterior, dividimos el intervalo $[0,1]$ *en* $n+1$ *subintervalos iguales con nodos igualmente espaciados.*

$$\frac{j}{n+1}, \quad j = 0, 1, \ldots, n+1,$$

terminamos con un funcional de costo

$$T(X) = \frac{1}{n+1} \sum_{j=0}^{n} \sqrt{1 + (n+1)^2 (x_{j+1} - x_j)^2}.$$

Las restricciones son $x_0 = x_{n+1} = 0$ *y*

$$\frac{1}{3} = \sum_{j=0}^{n} \frac{x_{j+1} + x_j}{2(n+1)}.$$

Esta ultima restricción se puede escribir de la forma

$$\frac{n+1}{3} = \sum_{j=1}^{n} x_j,$$

donde hemos tenido en cuenta las restricciones de frontera $x_0 = x_{n+1} = 0$. *En resumen. nos gustaria encontrar la solución óptima de*

$$\textit{Minimizar} \quad \sum_{j=0}^{n} \sqrt{1 + (n+1)^2 (x_{j+1} - x_j)^2}$$

sujeto a

$$x_0 = x_{n+1} = 0, \quad \frac{n+1}{3} = \sum_{j=1}^{n} x_j,$$

un problema de programación no lineal con restricciones lineales. La figura 5.15 muestra la aproximación resultante para un valor particular de n. *Note que hemos ignorado un multiplicador positivo del funcional.*

5.10. Ejercicios

1. Determine las geodesicas en una esfera. Intente adivinar cuales pueden ser las geodesicas, y argumente que estas son en realidad aquellas que dan el valor mínimo del funcional apropiado.

2. Investigue el ejercicio 9 del capítulo 1. intentando encontrar caminos óptimos.

3. Escriba la ecuacion de E-L para los siguientes funcionales:

$$I(u) = \int \left[(u')^2 + e^u \right] dx, \quad I(u) = \int u u' dx, \quad I(u) = \int x^2 (u')^2 dx.$$

4. Encuentre la solución óptima al problema

$$\text{mín}\left\{\int_0^1 [x'(t)^2 + 2x(t)^2]e^t dx : x(0) = 0, x(1) = e - e^2\right\}$$

y el valor número de este mínimo.

5. Resuelva el siguiente problema variacional

$$\text{Minimizar} \quad I(u) = \int_0^1 (u'(x)^2 + u(x)u'(x) + u(x)^2)dx$$

entre todas aquellas que satisfagan $u(0) = 0$, $u(1) = 1$

6. Resuelva el problema de minimizar el funcional

$$I(u) = \int_0^1 u'(x)^2, \quad u(0) = a_0, \quad u(1) = a_1,$$

entre todas aquellas funciones que satisfagan

$$0 = \int_0^1 u(x)\cos(b_i x)dx, \quad i = 1, 2, \ldots, N.$$

donde a_0, a_1 y b_i son parametros fijos.

7. Encuentre la función $u(x)$ que minimiza

$$I(u) = \int_0^1 (1 + u''(x)^2)dx$$

bajo las restricciones $u(0) = 0$, $u(1) = 1$, $u'(0) = 1$, $u'(1) = 1$.

8. Resuelva el problema anterior con el funcional

$$I(u) = \int_0^{\pi/2} (u''(x)^2 - u(x)^2 + x^2)dx,$$

sujeto a

$$u(0) = 1, \quad u(\pi/2) = 0, \quad u'(0) = 0, u'(\pi/2) = 1.$$

9. El principio de máxima entropia selecciona la distribución de probabilidad sobre el semi intervalo $(0, \infty)$ que maximice la integral

$$H = -\int_0^\infty u(t)\log u(t)dt.$$

Si imponemos las restricciones

$$\int_0^\infty u(t)dt = 1, \quad \int_0^\infty tu(t)dt = 1/a,$$

compruebe que la función de distribucion mas probable es $u(t) = ae^{-at}$

10. Considere la función

$$f(y) = \text{mín}\left\{\int_0^{\log 2} [x'(t)^2 + x(t)^2]dx : x(0) = 1, x(\log 2) = y\right\}.$$

Encuentre una expresión explicita para $f(y)$ y usela para determinar la solución del problema

$$\text{mín}\left\{\int_0^{\log 2} [x'(t)^2 + x(t)^2]dx : x(0) = 1\right\},$$

Resuelva directamente este último problema y compare sus resultados.

11. Determine el mínimo valor de las integrales

$$\int_0^{\log 2} \left[(x'(t) - 1)^2 + x(t)^2\right] dx$$

entre todas las funciones $x(t)$.

12. Resuelva el problema de la cuerda de sección transversal variable del capítulo 1

a) Muestre que con el cambio

$$b(x) = W + \rho g \int_x^L a(s)ds,$$

el problema puede ser reformulado como

$$\text{Minimizar} \quad \left(-\frac{\rho g}{E}\right) \int_0^L \frac{b(x)}{b'(x)} dx$$

sujeto a

$$b(0) = W + \rho g V, \quad b(L) = W$$

b) resuelva el problema en esta forma, e interprete el resultado final en terminos de la formulación inicial.

13. En ocasiones, la formulación exacta de un problema variacional es imposible de resolver, y se deben hacer aproximaciones razonables. El problema del solido moviendose en un fluido descrito en el capítulo 1 es un ejemplo. Haga simplificaciones razonables como en el ejemplo 5.24 para obtener una buena aproximación al perfil del solido en movimiento.

14. Estudie el problema

$$\text{Minimizar} \quad \frac{1}{2} \int_0^1 u'(x)^2 dx$$

sujeto a

$$\int_0^1 1u(x)^2 dx \leq 2, \quad \int_0^1 u(x)dx = 1.$$

15. (En problema de obstaculo elemental) Trate de entender el siguiente problema variacional

$$\text{Minimizar} \quad \int_0^1 u'(x)^2 dx$$

sujeto a

$$u(0) = u(1) = 0, \quad u(x) \geq u_0(x),$$

donde u_0 es una función dada de tal manera que $u_0(0)$ y $u_0(1)$ son negativas pero u_0 es positiva en algun punto en el intervalo $(0, 1)$ cúando

$$u_0(x) = -\frac{1}{8} - \frac{1}{2}x^2 + \frac{1}{2}x$$

y cúando

$$u_0(x) = \frac{1}{4}\sin(10x) - \frac{1}{2} - 4x^2 + 4x.$$

16. El ejercicio 12 del capítulo 1 se puede entender y resolver como un problema variacional bajo restricciones puntuales. Introduzca un multiplicador (una función) e intente determinar las soluciones óptimas.

17. Encuentre una solución aproximada al problema de braquioscrona

$$I(u) = \int_0^1 \frac{\sqrt{1 + u'(x)^2}}{\sqrt{x}} dx,$$
$$u(0) = 0, \quad u(1) = -1,$$

como se describe en la sección 5.9. Haga la misma cosa para el problema del cable colgante

$$\text{Minimizar} \quad \int_0^1 u(x)\sqrt{1 + u'(x)^2} dx$$

sujeto a

$$u(0) = u(1) = 0, \quad \frac{3}{2} = \int_0^1 \sqrt{1 + u'(x)^2} dx.$$

Capítulo 6

Control Óptimo

6.1. Introducción

El control óptimo es un tema muy importante en optimización, con muchas aplicaciones en diferentes áreas, especialmente en ingeniería. En este último capítulo, simplemente estudiaremos las ideas básicas para abordar estos problemas. En particular, nos enfocaremos en el principio de Pontryagin del máximo, tratando de insistir en su importancia a traves de varios ejemplos.

El formato usual de un problema de control óptimo es el siguiente. El estado de un cierto sistema esta descrito por un número de parametros.

$$x = (x_1, x_2, \ldots, x_n),$$

el cual evoluciona de acuerdo a la ecuación de estado

$$x'(t) = f(t, x(t), u(t)),$$

donde

$$u = (u_1, u_2, \ldots, u_m)$$

representa el control ejercido en el sistema (con el objetivo en mente de controlarlo). Este vector de control normalmente debería satisfacer varios tipos de restricciones dependiendo la naturaleza del problema. Consideraremos solamente la restricción $u(t) \in K \subset \mathbb{R}^m$ para todos los t, donde K está dado a prori. La ecuación de estado esta complementada con condiciones iniciales y/o finales.

$$x(0) = x_0 \quad x(T) = x_T,$$

donde T es el horizonte de tiempo que estamos considerando. Tambien debemos tener un funcional de costo que mida qué tan bueno es un control dado u. La forma de este funcional objetivo es

$$I(x, u) = \int_0^T F(t, x(t), u(t))dt,$$

donde

$$F : (0,T) \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$$

es un integrando conocido asociado con el costo que deseamos medir. Una pareja (x,u) se dice que es admisible o factible si cumple

1. Las restricciones en el control: $u(t) \in K$ para todos los $t \in (0,T)$.
2. La ley de estado: $x'(t) = f(t,x(t),u(t))$ para todos los $t \in (0,T)$.
3. La condiciones en los puntos extremos: $x(0) = x_0$, $x(T) = x_T$.

Como se menciono anteriormente, las condiciones en los puntos extremos pueden variar de tener ninguna condicion en los puntos a tenerlas en los dos. El problema de control óptimo consiste en encontrar (determinando o aproximando) un par admisible (X,U) tal que

$$I(X,U) \leq I(x,u)$$

para todos los otros pares factibles (x,u).

Note que este tipo de problemas incluye como un caso muy particular, los problemas variacionales en los cuales tratamos de minimizar

$$I(x) = \int_0^T F(t,x(t),x'(t))dt$$

entre todos los campos $x(t)$ con $x(0) = x_0$, $x(T) = x_T$. De hecho este problema es equivalente a

$$x'(t) = u(t), \quad x(0) = x_0, \quad x(T) = x_T,$$
$$I(x,u) = \int_0^T F(t,x(t),u(t))dt.$$

En este capítulo nos enfocaremos esencialmente en las condiciones necesarias de optimalidad. Estamos buscando condiciones, en forma diferencial, que las soluciones óptimas de los problemas de control deben satisfacer. Daremos por sentado que nuestros problemas siempre tienen la regularidad necesaria para escribir tales condiciones de optimalidad. Todo un campo fundamental (optimización no regular) estudia todas aquellas situaciones en las cuales la regularidad no puede asumirse, y de hecho, es uno de los grandes temas para entender. El Análisis no regular se deja para textos mas avanzados. Vease [9].

6.2. Multiplicadores y el Hamiltoniano

Primero trataremos el caso en el cual el conjunto K es todo $\mathbb{R}^n$, asi que no hay restricciones de cúales valores pueden tomar los controles admisibles. Por lo tanto estamos interesados en minimizar

$$I(x,u) = \int_0^T F(t,x(t),u(t))dt$$

entre todos los pares (x, u) tales que

$$x'(t) = f(t, x(t), u(t))$$

junto con las condiciones apropiadas en los puntos extremos. Como ya hemos notado, la ecuación de estado puede ser considerada como una restricción puntual que puede ser tratada introducuendo un multiplicador o coestado $p(t)$. Por lo tanto, consideremos el funcional aumentado

$$I^*(x, u, p, x') = \int_0^t [F(t, x(t), u(t)) + p(t)(f(t, x(t), u(t)) - x'(t))]dt.$$

De nuestra experiencia con restricciones y multiplicadores, parece plausible que las soluciones óptimas para nuestro problema inicial de contol óptimo debe ser una solución de la ecuación E-L para I^* visto como una función de las variables (x, u, p, x'). Dado que el principio de Pontryagin del máximo es un enunciado mas general que este y sera tratado mas adelante, no insistitemos en la validad de esta afirmación en este punto. Si escribimos

$$G(t, u, p, x, u', p', x') = F(t, x, u) + p(f(t, x, u) - x'),$$

entonces el sistema E-L puede ser escrito como

$$\frac{d}{dx}\left[\frac{\partial G}{\partial x'}\right] = \frac{\partial G}{\partial x}, \quad \frac{d}{dt}\left[\frac{\partial G}{\partial u'}\right] = \frac{\partial G}{\partial u}, \quad \frac{d}{dt}\left[\frac{\partial G}{\partial p'}\right] = \frac{\partial G}{\partial p},$$

Esto es,

$$\begin{aligned} 0 &= \frac{\partial F}{\partial x}(t, x, u) + p\frac{\partial f}{\partial x}(t, x, u) + p', \\ 0 &= \frac{\partial F}{\partial u}(x, u, t) + p\frac{\partial f}{\partial u}(t, x, u), \\ 0 &= x' - f(x, u, t). \end{aligned}$$

Definiendo el Hamiltoniano del problema como $H = F + pf$, estas ecuaciones se puden escribir como

$$p' = -\frac{\partial H}{\partial x}, \quad \frac{\partial H}{\partial u} = 0, \quad x' = f(t, x, u).$$

Debemos completar este sistema con las condiciones en los extremos. Note qie solo dos derivadad (p' y x') aparecen, asi que para determinar la solución necesitamos dos condiciones. Estas son las restricciones en los extremos para el estado x completadas con las condiciones de transversalidad para el multiplicador p de acierdo con la siguiente regla:

Afirmación de transversalidad 6.1 *Si en un punto extremos dado (inicial o final) tenemos una condición en el estado, no imponemos la condición de transversalidad correspondiente, pero si el estado es libre, entonces la condición de transversalidad $p = 0$ en el punto extremo dado debe tomarse.*

Como ya hemos argumentado en otras situaciones, las soluciones óptimas deben buscarse entre las soluciones de estas condiciones de optimalidad. Veamos varios ejemplos.

Ejemplo 6.2 *Si*

$$I(x,u) = \int_0^1 u(t)^2 dt, \quad x'(t) = u(t) + ax(t),$$

donde $a \in \mathbb{R}$ es una constante, nos gustaria determinar el control óptimo bajo la condición inicial $x(0) = 1$. En este ejemplo el Hamiltoniano es

$$H = u^2 + p(u + ax),$$

y la hip'ótesis que garantizan que las soluciones óptimas son precisamente las soluciones a las condiciones de optimalidad estan satisfechas. Esto sera justificado despues.

Manipulando estas condiciones de optimalidad llegamos a

$$u = -p/2, \quad p' = -ap, \quad x' = -p/2 + ax,$$

junto con $x(0) = 1$, $p(1) = 0$. Esta ultima condición es la condición de transversalidad en $t = 1$, dado que no tenemos la condición final para el estado (volveremos a este punto despues). Resolviendo para p obtenemos

$$p(t) = 0,$$

y entonces

$$x(t) = e^{at}$$

El control óptimo es entonces $u = 0$, como podriamos haber anticipado. en relidad esta situación no tiene mucho interes ya que no hay demanda en el sistema.

Suponga ahora que nos gustaria $x(0) = 1$ y $x(1) = 0$. La estrategia óptima ya no es $u = 0$, ya que este control lleva al sistema a $x = 1 = e^a$. Ya que ahora no tenemos tennemos condicion de transversalidad para p y

$$p(t) = ce^{-at}$$

Tomando esta expresión en la ecuación de estado obtenemos

$$x(t) = \frac{c}{4a}e^{-at} + de^{at}$$

Las constantes c y d se deben escoger de tal manera que las condiciones en los puntos extremos se satisfagan. Esto lleva a un sistema lineal cuya solución es

$$c = \frac{2ae^a}{\sinh a}, \quad d = -\frac{e^{-a}}{2\sinh a}$$

En esta ocasión el control óptimo es

$$u(t) = -\frac{ae^{a(1-t)}}{\sinh a}.$$

Ejemplo 6.3 *El costo es el mismo que en el ejemplo anterior*

$$I(x,u) = \int_0^1 u(t)^2 dt,$$

pero la ecuación de estado es

$$x''(t) = u(t)$$

con condiciones en los extremos

$$x(0) = x'(0) = 1, \quad x(1) = 0.$$

El control u es el acelerador, y el costo es una medida del gasto de combustible. Primero reducimos la ecuación de segundo orden a una de primer orden con componentes $x_1 = x$, $x_2 = x'$, asi

$$x_1' = x_2, \quad x_2' = u,$$

y condiciones en los extremos

$$x_1(0) = 1, \quad x_1(1) = 0, \quad x_2(0) = 1.$$

El Hamiltoniano es

$$H = u^2 + p_1 x_2 + p_2 u,$$

y las ecuaciones de optimalidad junto con las condiciones en los puntos extremos son

$$\begin{aligned} u &= -p_2/2 \\ p_1' = 0, \quad p_2' &= -p1, \\ x_1' = x_2, \quad x_2' &= u, \\ x_1(0) = 1, \quad x_1(1) &= 0, \\ x_2(0) = 1, \quad p_2(1) &= 0. \end{aligned}$$

Con esto no es dificil encontrar la solución óptima

$$\begin{aligned} u(t) &= 6(t-1) \\ x_1(t) = t^3 - 3t^2 + t + 1 x_2(t) &= 3t^2 - 6t + 1 \end{aligned}$$

Estas condiciones son necesarias para una solución óptima, pero pueden no ser suficientes en el sentido que pude haber otro tipo de soluciones que no son óptimas. A esta altura, no sorprendera al lector si aseguramos que la convexidas esta involucrada en tratar de establecer que las condiciones necesarias de optimalidad tambien son suficientes. De hecho, podemos probar el siguiente resultado. Nos apegamos a la notación inicial para un problema de control óptimo

$$\begin{aligned} I(x,u) = \int_0^T F(t,x(t),u(t))dt, \\ x'(t) = f(t,x(t),u(t)), \\ (x(0) = x_0), \quad (x(T) = x_t), \end{aligned}$$

Hemos puesto las condiciones en los puntos extremos entre parentesis para indicar que estas pueden o no estar presentes.

Teorema 6.4 *Sea f lineal en (x,u) y F convexo en (x,u) para cada t fija. Entonces todas las soluciones con del sistema de optimalidad con las condiciones apropiadas en los puntos extremos (incluyendo optimalidad) sera una solución óptima del problema de control.*

Suponga que el par (x,u) satisface todas las condiciones de optimalidad, y sea $(\tilde{x},\tilde{u})$ cualquier otro par admisible. Mediremos la diferencia

$$I(\tilde{x},\tilde{u}) - I(x,u)$$

y concluiremos que no puede er negativa. Esto implica que (x,u) es de hecho óptima.

Debido a las hipótesis de linealidad y convexidad supuestas en el enunciado del teorema podemos escribir

$$\begin{aligned}
&I(\tilde{x},\tilde{u}) - I(x,u) = \int_0^T [F(t,\tilde{x},\tilde{u}) - F(t,x,u)]dt \\
&\geq \int_0^T \left[\frac{\partial F}{\partial x}(t,x,u)(\tilde{x}-x) + \frac{\partial F}{\partial u}(t,x,u)(\tilde{u}-u)\right] dt \\
&= \int_0^T \left[-p\frac{\partial f}{\partial x}(t,x,u) - p')(\tilde{x}-x) - p\frac{\partial f}{\partial u}(t,x,u)(\tilde{u}-u)\right] dt \\
&= -\int_0^T p\left[(x'-\tilde{x}) + \frac{\partial f}{\partial x}(t,x,u)(\tilde{x}-x) + \frac{\partial f}{\partial u}(t,x,u)(\tilde{u}-u)\right] dt \\
&= -\int_0^T p\left[f(t,x,u) - f(t,\tilde{x},\tilde{u}) + \frac{\partial f}{\partial x}(t,x,u)(\tilde{x}-x) + \frac{\partial f}{\partial u}(t,x,u)(\tilde{u}-u)\right] dt \\
&= 0
\end{aligned}$$

Notese como las condiciones en los puntos extremos y la transversalidad se usan para evaluar las contribuciones de los puntos extremos se anulan en las integraciones por partes.

En ocasiones, se desea forzar restricciones integrales adicionales.

Ejemplo 6.5 *Suponga que un proceso particular esta descrito por una función $x(t)$ comenzando en $x(0)=1$. Nuestro deseo es llevar al sistema a $x(T)=0$ en el menor tiempo posible, donde el sistema evoluciona de acuerdo a la ley*

$$x'(t) = ax(t) + u(t),$$

donde a es una constante dada, y debemos respetar la restricción

$$K = \int_0^T u(t)^2 dt,$$

para una constante fija $K>0$. Es precisamente esta restricción integral la que hace al problema interesante. porque de lo contrario, T podria ser tan pequeño

como quisieramos. Despues de nuestra experiencia con restricciones integrales en el capítulo anterior, vamos a introducir un multiplicador $s \in \mathbb{R}$ asociado con esta restricción, y escribimos

$$H = 1 + su^2 + p(ax + u)$$

para el Hamiltoniano, donde s se determinara luego. Es mas, note que podemos escribir

$$H = s\left(\frac{1}{s} + u^2 + \frac{p}{s}(ax + u)\right),$$

y reemplasando p por p/s, concluimos que las condiciones de optimalidad son ls mismas para el Hamiltoniano

$$H = 1 + u^2 + p(ax + u),$$

pero en este caso no hay constante adicional s. Esta ha sido incorporada de alguna manera en el co estado p. Las condiciones de optimalidad son

$$p' = -ap, \quad 2u + p = 0, \quad x' = ax + u.$$

Resolviendo para u y p y reemplazandolas en la ecuación de estado, obtenemos

$$x' = ax + ce^{-at}$$

para cierta constante c. Resolviendo esta última ecuación, llegamos a

$$x(t) = de^{at} + \frac{c}{2a}e^{-at}.$$

Imponemos las tres condiciones

$$x(0) = 1, \quad x(T) = 0, \quad K = \int_0^T u(t)^2 dt,$$

y tenemos

$$\begin{aligned} 1 &= d + \frac{c}{2a}, \\ 0 &= de^{aT} + \frac{c}{2a}e^{-aT}, \\ K &= \frac{c^2}{8a}(1 - e{-}2aT). \end{aligned}$$

Manipulando estas ecuaciones obtenemos un T minimo,

$$T = -\frac{1}{2a}\log\left(1 - \frac{a}{2K}\right)$$

y como control óptimo

$$u(t) = -2Ke^{-at},$$

y el estado óptimo asociado

$$x(t) = e^{at} - \frac{4K}{a}\operatorname{sinh}(at).$$

Mote que la formula para T requiere que $a < 2K$. Si $a = 2K$, entonces $T = +\infty$ y el sistema tiendea 0 a medida que $t \to \infty$ pero nunca lo alcanza. Si $a > 2K$, el sistema ni siquiera se acerca al estado 0, dado que $x(t)$ tiende a $+\infty$.

Todos los ejemplos que hemos visto hasta ahora cumplen las hipótesis del teorema 6.4, asi que las soluciones que hemos encontrado son de hecho óptimas.

6.3. El principio de Pontryagin

Si deseamos acercarnos a hipótesis mas realistas para problemas de control óptimo, debemos preocuparns acerca de la posiblidad de tener ciertas restricciones en el tamaño de los controles admisibles. Esto es razonable, dado que a priori, un sitema dado puede no soportar el efecto de un control de tamaño arbitrario, ya sea porque este lo haría colapsar, o porque lo sacaria del regimen en el cual la ley de estado es valida, o simplemente porque no tenemos los medios para aplicar un control de tamaño arbitrario. En cualquier caso, debemos considerar una restricción en el tamaño de los controles factibes. Tipicamente, esto es formulado exigiendo

$$u(t) \in K$$

donde K es un subconjunto apropiado del espaso en el cual los controles toman valores. Por lo tanto en esta sección estamos interesados con el problema

$$\text{Minimizar} \quad \int_p^T F(t, x(t), u(t))dt$$

sujeto a

$$x'(t) = f(t, x(t), u(t)),$$
$$u(t) \in K, \quad x(0) = x_0, \quad (x(T) = x_T).$$

Estamos especialmete interesados en entender las condiciones necesarias de optimalidad que deben satisfacer las soluciones óptimas.

Si recordamos la discución relativa a la ecuación de Bellman en el capítulo anterior, podemos considerar la siguiente función valor

$$S(t,x) = \min_u \left\{ \int_t^T F(\tau, y(\tau), u(\tau))d\tau : y'(\tau) = f(\tau, y(\tau), u(\tau)), \right.$$
$$\left. u(\tau) \in K, y(t) = x \right\}.$$

La propiedad fundamental que una función valor debe satisfacer para $t' > t$ es

$$0 = \min_y \left\{ \min_v \left\{ \int_t^{t'} F(\tau, z(\tau), v(\tau))d\tau : z'(\tau) = f(\tau, z(\tau), v(\tau)), v(\tau) \in K, \right.\right.$$
$$\left.\left. z(t) = x, z(t') = x + y(t' - t) \right\} + S(t', x + y(t' - t)) \right\}$$

Esta condicion se puede reescribir como

$$S(t,x) = \min_y \left\{ \min_v \left\{ \frac{1}{t'-t} \int_t^{t'} F(\tau, z(\tau), v(\tau))d\tau : z'(\tau) = f(\tau, z(\tau), v(\tau)), v(\tau) \in K, \right.\right.$$
$$\left.\left. z(t) = x, z(t') = x + y(t' - t) \right\} + \frac{S(t', x + y(t' - t)) - S(t,x)}{t' - t} \right\}.$$

Tomando limites cuando $t' \searrow t$, concluimos (¿por qué?) que

$$0 = \min_y \left\{ \min_{v \in K} \{F(t,x,v) : y = f(t,x,v)\} + \frac{\partial S}{\partial t}(t,x) + y \frac{\partial S}{\partial x}(t,x) \right\},$$

que puede reorganizarse como

$$\frac{\partial S}{\partial t}(t,x) = - \min_{v \in K} \left[F(t,x,v) + f(t,x,v) \frac{\partial S}{\partial x}(t,x) \right]. \tag{6.1}$$

La pregunta es ¿qué tipo de informacion nos da esta ecuación acerca del par óptimo (x(t),u(t))? El hecho que (x(t),u(t)) son óptimas significa que

$$S(t, x(t)) = \int_t^T F(\tau, x(\tau), u(\tau))d\tau, \quad x'(\tau) = f(\tau, x(\tau), u(\tau)),$$

para cada tiempo t. Si derivamos con respecto a t y usamos 6.1, podemos escribir

$$\begin{aligned} -F(t, x(t), u(t)) &= \frac{\partial S}{\partial t}(t, x(t)) + x'(t) \frac{\partial S}{\partial x}(t, x(t)) \\ &= - \min_{v \in K} \left[F(t, x(t), v) + f(t, x(t), v) \frac{\partial S}{\partial x}(t, x(t)) \right] \\ &\quad + f(t, x(t), u(t)) \frac{\partial S}{\partial x}(t, x(t)), \end{aligned}$$

y por tanto

$$\begin{aligned} F(t, x(t), u(t)) + f(t, x(t), u(t)) &\frac{\partial S}{\partial x}(t, x(t)) \\ &= \min_{v \in K} \left[F(t, x(t), v) + f(t, x(t), v) \frac{\partial S}{\partial x}(t, x(t)) \right], \end{aligned}$$

por otra parte si definimos $p(t,x) = \frac{\partial S}{\partial x}(t,x)$, en el par óptimo obtenemos usando de nuevo

$$\frac{\partial S}{\partial t}(t,x(t)) = x'(t)\frac{\partial S}{\partial x}(t,x(t)) + F(t,x(t),u(t)),$$

que

$$\begin{aligned}\frac{d}{dt}p(t,x(t)) &= \frac{\partial^2 S}{\partial x \partial t}(t,x(t)) + f(t,x(t),u(t))\frac{\partial^2 S}{\partial x^2}(t,x(t)) \\ &= -\frac{\partial}{\partial x}\left(F(t,x(t),u(t)) + f(t,x(t),u(t))\frac{\partial S}{\partial x}(t,x(t))\right) \\ &+ f(t,x(t),u(t))\frac{\partial^2 S}{\partial x^2}(t,x(t)) \\ &= -\frac{\partial F}{\partial x}(t,x(t),u(t)) - p(t,x(t))\frac{\partial f}{\partial x}(t,x(t),u(t)).\end{aligned}$$

Por medio del Hamiltoniano

$$H(t,x,u,p) = F(t,x,u) + pf(t,x,u)$$

y la función $p(t) = p(t,x(t))$, podemos resumir las conclusiones anteriores en el siguiente enunciado. Note la relacion entre el multiplicador (co estado) $p(t)$ y la función valor $S(t,x)$:

$$p(t) = p(t,x(t)) = \frac{\partial S}{\partial x}(t,x(t))$$

si $x(t)$ es óptima

Teorema 6.6 *Principio de Pontryagin: Condiciones Necesarias. Si la pareja* $(x(t),u(t))$ *es óptima para nuestro problema de control original, debe existir una función* $p(t)$ *que cumple las siguientes condiciones:*

$$\begin{aligned}p'(t) &= -\frac{\partial H}{\partial x}(t,x(t),u(t),p(t)), \quad (p(T)=0), \\ H(t,x(t),u(t),p(t)) &= \min_{v\in K} H(t,x(t),v,p(t)), \\ x'(t) &= f(t,x(t),u(t)), \quad x(0)=x_0, \quad (x(T)=x_T).\end{aligned}$$

Note como la segunda condición en el enunciado anterior esta formulada como un PPNL para v dependiendo en varios parametros. Hemos escrito la condición de transversalidad y la condición final entre parentesis para enfatizar que una de las dos, pero no las dos al tiempo deben forzarse. El unico comentario en este encunciado tiene que ver con la condición de transversalidad $P(T)=0$. Si tenemos en mente la definición de $p(t)$ como

$$p(t) = \frac{\partial S}{\partial x}(t,x(t)),$$

la condicion de transversalidad significa que

$$\frac{\partial S}{\partial x}(t,x(T)) = 0.$$

Esta restricciñ refleja nada mes que el hecho que si el valor en el punto extremo derecho $x(T)$ es libre, la función valor debe alcanzar su mínimo cuando toma el valor del estado óptimo $x(T)$. En consecuencia la derivada debe anularse. Si en la formulación original tenemos una restricción fija en el estado final, $x(T) = x_T$, entonces esta condición reemplaza transversalidad. Lo mismo aplica para el tiempo inicial $t = 0$.

Tambien es interesante señalar que las condiciones en el resultado anterior son una generalizacion de aquellas en la sección 6.2 dado que si K es todo $\mathbb{R}^m$, entonces el mínimo del Hamiltoniano con respecto a v se alcanza cuando la derivada con respecto a v se anula.

Antes de estudiar las condiciones suficientes de optimalidad, veamos varios ejemplos para examinar el principio de Pontryagin.

Una de las familias mas comunes de problemas de control óptimo es aquella que estudia como ralizar una tarea en el mínimo tiempo posible. En esos casos, el funcional objetivo a ser minimizado es aquel del tiempo empleado en hacer la tarea.

Ejemplo 6.7 *Imagine para empezar, un objeto movil cuyo movimiento podemos controlar con su acelerador u, donde ela máxima aceleración posible es b, y el maximo poder de frenos es $-a$ i.e $-a \leq u \leq b$. De acuerdo a nuestra discución anterior, $K = [-a, b]$. Comenzando en reposo y terminando en reposo, nos gustaria viajar una distancia α en un tiempo mínimo. Cúal es la estrategia óptima para usar el acelerador?. El funcional de costo es*

$$I = \int_0 T1dt,$$

bajo las restricciones

$$\begin{aligned} x''(t) = u(t), \quad u \in K \\ x(0) = x'(0) = 0, \quad x(T) = \alpha, x'(T) = 0. \end{aligned}$$

Si transformamos esta ecuación de segundo orden en un sistema de primer orden enm la forma estandar por medio del cambio

$$x_1(t) = x(t), \quad x_2(t) = x'(t),$$

obtenemos

$$\begin{aligned} x_1' = x_2, \quad x_2' = u, \quad u \in K \\ x_1(0) = x_2(0) = 0, \quad x_1(T) = \alpha, \quad x_2(T) = 0. \end{aligned}$$

El Hamiltoniano del sistema sera

$$H(t, x, u, p) = 1 + p_1 x_2 + p_2 u,$$

y las condiciones necesarias de optimalidad de acuerdo con el principio de Pontryagin seran

$$p_1' = 0, \quad p_2' = -p_1.$$

Dado que $p_1 = -d$ es una constante, y $p_2 = dt + c$, donde c, d son constantes. Si tomamos esta informacion a la condición del mínimo, obtenemos que el control óptimo $u(t)$ debe minimizar el HamiltonianoH sobre K. En nuestra situación,

$$(dt + c)u(t) = \min_{-a \le v \le b} (dt + c)v.$$

Dado que la expresión $(dt + c)v$ es lineal en v, el mínimo anterir se alcanza en alguno de los extremos del intervalo $[-a, b]$ dependiendo en el signo de $(dt + c)$. Por lo tanto el control óptimo tendra la forma

$$u(t) = \begin{cases} -a, & dt + c > 0 \\ b, & dt + c < 0 \\ cualquier\ valor, & dt + c = 0. \end{cases}$$

Debido a la interpretación fisica del problema, tenemos $u(t) = b$ en el inicio, ya que de lo contrario nuestro vehiculo no se movera, y obviamanete mas adelante tendremos que usar el freno. Note tambien que en forma anterior para el control óptimo descarta la posibilidad de tener varios cambis entre aceleraciones positivas y negativas, ya que $dt + c$, siendo lineal en t puede pasar por cero a lo mas una vez. En conseciencia en control óptimo tendra la forma

$$u(t) = \begin{cases} b, & t \le t_0, \\ -a, & t \ge t_0 \end{cases}$$

El instante t_0 debe ser determinado en terminos de a, b y α. De hecho, tendremos que resolver

$$x''(t) = b, \quad x(0) = x'(0) = 0$$

con solución

$$x(t) = bt^2/2.$$

En un tiempo t_0 (a ser determinado) hay un cambio en la dinamica del sistema y debemos resolver

$$x''(t) = -a, \quad x(t_0)nt_0^2/2, \quad x'(t_0) = bt_0,$$

pidiendo, para $x(T) = \alpha$ y $x'(T) = 0$. Por lo que

$$x(t) = bt_0 \left(t - \frac{t_0}{2} \right) - a\frac{(t - t_0)^2}{2},$$

y las condiciones en el tiempo T llevan a un sistema de dos ecuaciones en las dos incognitas T y t_0. Despues de algunos calculos obtenemos

$$t_0 = \sqrt{\frac{2a\alpha}{b(a + b)}}, \quad T = \sqrt{\frac{2(a + b)\alpha}{ab}}$$

Ejemplo 6.8 *Nuestro siguiente ejemplo tambien es un problema de tiempo mínimo, asi que el funcional objetivo es el mismo, pero el estado del sistema es*

$$x_1' = -x_1 + u, \quad x_2' = u, \quad |u| \leq 1,$$

bajo condiciones iniciales arbitrarias

$$x_1(0) = a, \quad x_2(0) = b.$$

La tarea consiste en llevar el sistema al reposo,

$$x_1(T) = x_2(T) = 0$$

en el mínimo tiempo.

El Hamiltoniano es

$$H(tmxmump) = 1 + p_1(-x_1 + u) + p_2 u,$$

el cual, como en nuestro ejemplo anterior es lineal en u. Las ecuaciones para p_1 *y* p_2 *son*

$$p_1' = p_1, \quad p_2' = 0,$$

asi que $p_2 = d$ *y* $p_1 = ce^t$*, donde c y d son constantes. Dabido a la linalidad de H con respecto a u, el control óptimo tendra la forma*

$$u(t) = \begin{cases} -1 & p_1 + p_2 > 0, \\ 1, & p_1 + p_2 < 0, \\ \textit{cualquier valor}, & p_1 + p_2 = 0. \end{cases}$$

Debemos preguntarnos cuantas veces la expresión

$$p_1 + p_2 = d + ce^t$$

puede pasar a traves origen, con el objetivo de exhibir cuantos cambios de -1 *a* 1 *tendra el control óptimo y cuando, a traves de calculos aproximados encontrar la estrategia óptima. Dado que las condiciones iniciales son arbitrarias y no estan dadas explicitamente, la discución en terminos de formulas especificas que den el tiempo óptimo y los instantes en los que el control óptimo debe tomar lugar en terminos de a y b se vuelve una tarea tediosa y casi imposible de explicar, en estas ocaciones es mas enriquecedor analizar el problena por medio de las çurvas de cambio". Describiremos este tipo de análisis para el ejemplo que estamos trabajando.*

Como argumentamos anteriormente, el control óptimo hace que el proceso alterne entre las dinamicas de dos sistemas

$$\begin{cases} x_1' = -x_1 + 1, \\ x_2' = 1, \end{cases} \qquad \begin{cases} x_1' = -x_1 - 1 \\ x_2' = -1. \end{cases}$$

Las curvas integrales de estos sistemas comenzando en (a, b) *cuando* $t = 0$ *son, respectivamente,*

$$x_(t) = (a-1)e^{-t} + 1, \quad x_2(t) = t + b,$$

y

$$x_1(t) = (a+1)e^{-t} - 1, \quad x_2(t) = -t + b,$$

Que son sitematicamente dibujadas en la figura 6.1, donde la curva integral a traves del origen esta resaltada para cada caso

Si imaginamos estas dos familias de curvas en un solo digarama, la escogencia del control óptimo esta dada por determinar en un punto dado (a, b)*, siguiendo una de las correspondientes curvas integrales y, en un cierto instante, cambiar a la otra de tal manera que la segunda curva integral nos debe llevar al origen si es posible. Desde que solo una curva integral de cada sistema pasa a traves del origen (de hecho, desde que al tiempo no se le permite decrecer, estamos hablando de dos medias curvas), es cuestion de alcanzar una de estas dos medias curvas tan pronto como sea posible comenzando desde el punto inicial dado. Especificamente estas dos medias curvas son*

$$\Lambda_1 = \{(1 - e^{-s}, s) : s \leq 0\} \text{ para } u = 1$$
$$\Lambda_{-1} = \{(e^s - 1, s) : s \geq 0\} \text{ para } u = -1.$$

Sea Λ *la unión de estas dos curvas, sea* Λ^+ *la parte del plano sobre* Λ*, y* Λ^- *la parte del plando debajo de* Λ*. Es facil ver que si el punto* (a, b) *pertenece a* Λ^+*, debemos tomar* $u = -1$ *hasta que nos encontremos en* Λ_1*, cambiando en este instante a* $u = 1$*, dado que la curva nos llevara al origen. De la misma manera para condiciones iniciales* $(a, b) \in \Lambda^-$*, debemos escoger primero* $u = 1$ *hasta alcanzar* Λ_{-1}*, y entonces* $u = -1$ *dado que* Λ_{-1} *nos llevara al reposo. Estas son las estrategias óptimas, dado qeue cualquier otra manera de cambiar entre estas dos dinamicas nos haria "'gastar algo de tiempo"'. Vea la figura 6.2*

Tambien es interesante notar que el número de cambios en el control óptimo tambien puede ser determinado regreseando a la ecuación

$$p_1 + p_2 = d + ce^t$$

El numero de cambios corresponde a las raices de la ecuación

$$d + ce^t = 0$$

para constantes arbitrarias c, d*. Resolviendo para* t *obtenemos*

$$t = \log\left(-\frac{d}{c}\right).$$

Entonces conluimos que el control óptimo tendra a lo mas (cuando d/c *es negativo) un cambio.*

Ejemplo 6.9 *Un problema interesante que puede ser entendido como un análisis cualitatico usando el concepdo de cambiar de curva, es el control óptimo de un oscilador armonico, el cual es descrito por las ecuaciones.*

$$x_1' = x_2, \quad x_2' = -x_1 + u,$$

donde el control u esta restringido en tamaño $|u| \leq 1$. De nuevo nos gustaria saber como llevar el sistema al reposo comenzando en un estado inicial arbitrario (a_1, a_2).

El Hamiltoniano es

$$H(t, x, u, p) = 1 + p_1 x_2 + p_2(-x_1 + u).$$

Dado que es lineal en u, las estrategias óptimas siempre alternaran entre $u = 1$ y $u = -1$. Esas dos dinamicas estan representadas por las curvas integrales de los sistemas.

$$\begin{cases} x_1' = x_2 \\ x_2' = -x_1 + 1 \end{cases} \qquad \begin{cases} x_1' = x_2, \\ x_2' = -x_1 - 1. \end{cases}$$

no es dificil confirmas que las curvas integrales para el primer sistema son circulos concentricos centrados en $(1, 0)$, mientras que aquellas para el segundo son circulos concentricos centrados en $(-1, 0)$ (Figura 6.3)

El asunto es entener las estrategis óptimas para puntos iniciales arbitrarios en el plano. Para esto, examinamos las ecuaciones diferenciales para los coestados y vemos que informacion podemos extraer de ellos. Estos son

$$p_1' = p_2$$
$$p_2' = -p_1$$

Por diferenciacion tenemos

$$p_2'' + p_2 = 0,$$

asi que

$$p_2(t) = A\cos(t + B)$$

para constantes arbitrarias A, B. Sabemos que los cambios en el control óptimo estan dictados por los momentos en los cuales el co estado p_2 pasa por el origen.

Dada lla forma de p_2 concluimos que puede haber un numero arbitrario de cambios (dependiendo de las condiciones iniciales) y que estos pon producidos antes que π unidades de tiempo pasen. Note que las raices de $\cos(B + t)$ estan localizadas a π unidades unas de otras. Teniendo en mente esta información, y examinando primero aquellos puntos para los cuales no hay necesidad de cambiar en el contral, entonces aquellos que requieren un solo cambio, dos cambios y asi, es posible entender las estrategias óptimas. Estas conclusiones estan indicadas en la figura 6.4.

Ejemplo 6.10 *Desde un punto dado fijo, se lanza un proyectil con el objetivo de golpear un blanco localizado a* $3 + 5/6$ *unidades de distancia en* 3 *unidades de tiempo. Queremos alcanzar esto con el mínimo costo medido por la integral*

$$E = k \int_0^3 u(t)^2 dt, k > 0,$$

donde el control u *indica la aceleración del projectil. La ecuación de estado es* $x'' = u$*, y* u *esta restringido en signo y en tamaño:* $0 \leq u \leq 1$*. Las condiciones iniciales y finales son entonces*

$$x'(0) = x(0) = 0, \quad x(3) = 3 + \frac{5}{6}.$$

Transformando la ecuación de segundo orden a un sistema de primer orden, obtenemos inmediatamente el Hamiltoniano

$$H(t, x, u, p) = ku^2 + p_1 x_2 + p_2 u,$$

El cual es una función estrictamente convexa en u*. Las condiciones de optimalidad dicen que*

$$p_1' = 0, \quad p_2' = -p_1,$$

junto con la condicion de transversalidad $p_2(3) = 0$*, ya que el valor* $x_2(3)$ *es libre. En consecuencia, obtenemos* $p_2(t) = d(3 - t)$*, donde* $p_1 = d$ *es constante. La condición el el mínimo de* H *sobre* $u \in [0, 1]$*, la cual es cuadratica, lleva a estas tres situaciones, dependiendo de si el vertice de la parábola está en* $[0, 1]$ *o esta a la derecha o a la izquierda del intervalo. estos tres casos son (vease la figura 6.5):*

- $u = -p_2/2k = c(t - 3)$*,* $c = -d/3k$ *si* $c(t - 3) \in (0, 1)$*;*
- $u = 0$ *si* $c(t - 3) \leq 0$*;*
- $u = 1$ *si* $c(t - 3) \geq 1$*.*

Pero, desde que el control no puede ser identicamente cero (ya que el proyectil no se movería) solo son posibles dos casos, en los cuales c *siempre es no positiva (Figura 6.6):*

1. $0 < c(t - 3) < 1$*,* $0 \leq t < 3$*: En este caso debemos resolver el problema*

$$x'' = c(t - 3), \quad x(0) = x'(0) = 0, x(3) = 3 + \frac{5}{6}.$$

Despues de algunos cálculos obtenemos

$$x(t) = \frac{c}{6}t^3 - \frac{3c}{2}t^2, \quad c = -\frac{1}{3} - \frac{5}{54}.$$

Sin embargo, para este valor de c *del control óptimo* $u(t) = c(t-3)$ *siempre viola la restricción* $u \in [0, 1]$*, ya que* $u(0) > 1$*. Esta no puede ser la solución buscada.*

2. Existe $t_0 \in (0,1)$, de tal manera que $c(t_0 - 3) = 1$. En esta solución

$$u(t) = \begin{cases} 1, & t \leq t_0 \\ c(t-3), & t \geq t_0. \end{cases}$$

Resolvemos el problema

$$x'' = u, \quad x(0) = x'(0) = 0, \quad x(3) = 3 + \frac{5}{6},$$

en dos pasos. Primero

$$x''(t) = 1 \text{ para } t \leq t_0, \quad x(0) = x'(0) = 0$$

que tiene la solución $x(t) = t^2/2$. A continuación imponiendo la continuidad en t_0 para x y x',

$$x''(t) = c(t-3) \text{ para } t \geq t_0$$
$$x(t_0) = \frac{t_0^2}{2}, \quad x'(t_0), \quad x(3) = 3 + \frac{5}{6},$$

donde tenemos en cuenta que $c(t_0 - 3) = 1$. Despues de hacer todos estos cálculos, llegamos a una ecuación polinomial para t_0.

$$t_0^3 - 9t_0^2 + 23t_0 - 15 = 0, \quad t_0 \in (0,3).$$

El único valor admisible es $t_0 = 1$, y por lo tanto, $c = -1/2$. El control óptimo es

$$u(t) = \begin{cases} 1, & t \leq 1, \\ -\frac{t}{2} + 32, & t \geq 1. \end{cases}$$

Ejemplo 6.11 *Nuestro proximo ejemplo describe una situación el la cual en control tiene dos componentes, esto es para enfatizar las diferencias principales con los ejemplos examinados anteriormente.*

Suponga que un sistema obedece la ecuación de estado

$$x_1' = -x_2 + u_1, \quad x_2' = x_1 + u_2,$$

donde esta vez en control que ejercemos en el sistema debe respetar la restricción

$$u_1^2 + u_2^2 \leq 1$$

Nos guataria determinar el intervalo de tiempo mas corto en el cual podemos llevar el sistema al reposo desde una posición inicial arbitraria (a,b). El Hamiltoniano es

$$H(t,x,u.p) = 1 + p_1(-x_2 + u_1) + p_2(x_1 + u_2),$$

y la condición sobre el mínimo sobre el control (u_1, u_2) es

$$p_1u_1 + p_2u_2 = \text{mín}\{p_1v_1 + p_2v_2 : v_1^2 + v_2^2 \leq 1\}.$$

Este es un ejercicio simple de programación no lineal (Capítulo 3). En este punto nuestros lectores no tendran problema en encontrar la solución óptima

$$u_1 = -\frac{p_1}{\sqrt{p_1^2 + p_2^2}}, \quad u_2 = -\frac{p_2}{\sqrt{p_1^2 + p_2^2}},$$

el cual da el control óptimo una vez los multiplicadores (co estados) sean conocidos. Las ecuaciones para estos son

$$p_1' = -p_2, \quad p_2 = p_1.$$

Poniendo una en la otra llegamos a

$$p_1(t) = \rho_0 \cos(t + \theta_0), \quad p_2(t) = \rho_0 \sin(t + \theta_0),$$

donde hemos usado la forma $\rho_0 \cos(t+\theta_0)$ con ρ_0, θ_0 constantes arbitrarias, para la solución general de la ecuación $p_1'' + p_1 = 0$. De esta forma, el control óptimo es

$$u_1(t) = -\cos(t + \theta_0), \quad u_2(t) = -\sin_(t + \theta_0).$$

Ahora usamos el sistema de estado

$$x_1' = -x_2 - \cos(t + \theta_0), \quad x_2' = x_1 - \sin(t + \theta_0).$$

De nuevo, derivando y juntando las ecuaciones, obtenemos

$$x_1'' + x_1 = 2\sin(t + \theta_0).$$

La solución general de esta ecuación es

$$x_1(t) = -t\cos(t + \theta_0) + \rho_1 \cos(t + \theta_1),$$

y en consecuencia

$$x_2(t) = -t\sin(t + \theta_0) + \rho_1 \sin(t + \theta_1).$$

Finalmente, las condiciones iniciales y finales nos llevan a

$$\begin{aligned} a &= \rho_1 \cos\theta_1, \quad b = \rho_1 \sin\theta_1 \\ 0 &= -T\cos(T + \theta_0) + \rho_1 \cos(T + \theta_1), \\ 0 &= -T\sin(T + \theta_0) + \rho_1 \sin(T + \theta_1). \end{aligned}$$

De esto es inmediato obtener

$$T = \rho_1 = \sqrt{a^2 + b^2}, \quad \theta_0 = \theta_1,$$

y la estrategia óptima

$$u_1(t) = -\cos(t + \theta_0), \quad u_1(t) = -\sin(t + \theta_0),$$

donde θ_0 es el argumento inicial del estado inicial (a, b).

Una de las diferencias mas importantes que descubrimos en este ejemplo, en contraste de las situaciones en las cuales el control es simplemete un número, es qye la frontera de una region conexa de dos o mas variables no es disconexa, y por lo tanto, los controles óptimos no necesitan saltar abruptamente (aunque a veces lo hacen) de un punto a otro, pero los cambios en la dinamica del sistema toman lugar de forma suave.

Hemos estudiado varios ejemplos en los cuales haciendo uso de el principio de Pontryagin, hemos encontrado aparentemente las soluciones óptimas a los problemas. Como en otras situaciones que analizamos anteriormente, la pregunta es si podemos convencernos que estas soluciones calculadas son de hecho óptimas. Este es obviamente un punto importante. Es que si las condiciones necesarias de optimalidas son tambien suficientes. De nuevo, la convexidad juega un papel importante en esta discución.

Teorema 6.12 *Suficiencia de las condiciones de optimalidad Suponga que $F(t,x,u)$ es convexo en (x,u), y $f(t,x,u)$ es lineal en (x,u). Si la tipleta $(x(t),u(t),p(t))$ satisface*

$$\begin{aligned} p'(t) &= -\frac{\partial H}{\partial x}(t,x(t),u(t),p(t)), \quad (p(T)=0), \\ H(t,x(t),u(t),p(t)) &= \min_{v\in K} H(t,x(t),v,p(t)), \\ x'(t) &= f(t,x(t),u(t)), \quad x(0)=x_0, \quad (x(T)=x_T), \end{aligned}$$

donde

$$H(t,x,u,p) = F(t,x,u) + pf(t,x,u)$$

es el Hamiltoniano del sistema, y K es un conjunto convexo, entonces el par $(x(t),u(t))$ es una solución óptima del correspondiente problema de control óptimo.

Este resultado no es dificil de justificar despues de la experiencia que ya tenemos con convexidad. La condicion en el mínimo significa que

$$g(s) = H(t,x(t),u(t)+s(v-u(t)),p(t)), \quad s\in[0,1],$$

como una funcion de s para un t fijo y $v\in K$, tiene un mínimo en $s=0$. Note como el vector $sv+(1-s)u(t)$ pertenece a K si este conjunto es convexo. En esta situación lo unico que podemos asegurar es que $g'(0)\geq 0$ (mínimo de un lado). Por lo que,

$$0\leq g'(0) = \frac{\partial H}{\partial u}(t,x(t),u(t),p(t))(v-u(t)), \quad v\in K. \tag{6.2}$$

Si $(\tilde{x}(t),\tilde{u}(t))$ es cualquier otro par factible para nuestro problema de contol,

tenemos

$$\begin{aligned}
&I(\tilde{x},\tilde{u}) - I(x,u) \\
&\quad = \int_0^T [F(t,\tilde{x}(t),\tilde{u}(t)) + p(t)(f(t,\tilde{x}(t),\tilde{u}(t)) - \tilde{x}'(t)) \\
&\quad - F(t,x(t),u(t)) - p(t)(f(t,x(t),u(t)) - x'(t))]dt \\
&\quad = \int_0^T [H(t,\tilde{x}(t),\tilde{u}(t),p(t)) - H(t,x(t),u(t),p(t)) \\
&\quad - p(t)(\tilde{x}'(t) - x(t))]dt \\
&\quad \geq \int_0^T \left[\frac{\partial H}{\partial x}(t,x(t),u(t),p(t))(\tilde{x}(t) - x(t)) \right. \\
&\quad + \left. \frac{\partial H}{\partial u}(t,x(t),u(t),p(t))(\tilde{u}(t) - u(t)) + p'(t)(\tilde{x}(t) - x(t))\right]dt \\
&\quad \geq 0,
\end{aligned}$$

debido a la ecuación que satisface $p(t)$ y (6.2). Y por lo tanto el par $(x(t), u(t))$ es verdaderamente una verdadera solución óptima. Tambien hemos utilizado de manera escencial la convexidad de H con respecto de (x, u), la cual es garantizada debido a la convexidad de F y la linealidad de f. La presencia del multiplicador frente a f (y en particular su signo) nos evita aparentemente relajar la linealidad de f. Este hecho se deja a un curso mas avanzado en control óptimo..

Muy a menudo, no se pueden encontrar soluciones óptimas para problemas de control, para porque hay muchas (o muy pocas) variables involucradas, asi que es casi imposible manejarlas a mano; o porque las condiciones de optimalidad no se pueden resolver de manera explicita o es muy complicado y tedioso encontrar formulas explicitas.

Ejemplo 6.13 *Un objeto movil en el plano puede ser contro lado por dos parametros r_1 y r_2, expresando la rapidez con la cual la direccion del movimiento se pude cambiar (velocidad angular de movimiento) y el modulo de la valocidad respectivamente. Las ecuaciones de movimiento son*

$$x''(t) = \cos\theta(t)r_2(t), \quad y''(t) = \sin\theta(t)r_2(t), \theta'(t) = r_1(t).$$

Las restricciones en los pares factibles (r_1, r_2) se puden escribir normalmente exigiendo

$$(r_1, r_2) \in K,$$

donde K es el conjunto de controles admisibles. El objetivo es cambiar la posición del objeto de, digamos (a_0, b_0), a (a_1, b_1) en un tiempo mínimo. El sistema equivalente de primer orden es

$$\begin{aligned}
x_1' = x_2, \quad x_2' = \cos\theta r_2, \\
x_3' = x_4, \quad x_4' = \sin\theta r_2, \\
\theta' = r_1
\end{aligned}$$

y el Hamiltoniano es

$$H = 1 + p_1 x_2 + p_2 \cos\theta r_2 + p_3 x_4 + p_4 \sin\theta r_2 + p_5 r_1.$$

Las condiciones de optimalidad se escriben como

$$\begin{aligned} p_1' = 0, \quad p_2' &= -p_1, \\ p_3' = 0, \quad p_4' &= -p_3, \\ p_5' &= p_2 r_2 \sin\theta - p_4 r_2 \cos\theta, \\ x'' &= \cos\theta r_2, \\ y'' = \sin\theta r_2, \quad \theta' &= r_1, \end{aligned}$$

donde $r = (r_1, r_2)$ debe ser la solución óptima de

$$\min_{(r_1,r_2)\in K}((p_2\cos\theta p_4 \sin\theta)r_2 + p_5 r_1).$$

Incluso para escogencias simples para el conjunto K (como un rectangulo o una elipse)m es casi imposible determinar de manera explicita la solución óptima. Por otro lado, en esta situacion particular las restricciones en el estado (en adicionn de aquellas en los controles) en la forma de obstaculos a evadir son muy naturales. Esto, sin embargo esta mas alla de los objetivos de este texto.

Ejemplo 6.14 *La economia de un cierto pais sigue la ley*

$$k' = f(k) - (\lambda + \mu)k - c,$$

donde k es la razon de inversión por unidad de valor, f es la funcion de producción, μ y λ son parametros relacionados con la depreciación y el crecimiento laboral respectivament, y c es el consumo por unidad de valor. El objetivo de este pais es escoger el consumo para maximizar la integral de bienestar sobre un intervalo de tiempo dado

$$\int_0^T e^{-\delta t} u(t) dt$$

donde δ es el parametro de descuento en el tiempo y u es la función de utilidad. Esta ultima funcion debe satisfacer la ecuación

$$\eta = -\frac{cu''}{u'}$$

para una constante conocida η, a la cual se llama la elasticidad de la utilidad marginal. Si asumimos que k es conocido en ambos en el tiempo inicial y final y $u(T)$ es tambien conocido, nos gustaria decidir el consumo óptimo.

Si cambiamos la notación para hacer la formulación mas transparente, encontramos poniendo

$$x_1 = k, \quad x_2 = u, \quad x_3 = u', \quad v = c,$$

que el Hamiltoniano es

$$H = e^{-\delta t}x_2 + p_1(f(x_1) - (\lambda+\mu)x_1 - v) + p_2 x_3 - p_3\eta\frac{x_3}{v},$$

y las condiciones de optimalidad son

$$\begin{aligned}
x_1' &= f(x_1) - (\lambda+\mu)x_1 - v,\\
x_2' &= x_3,\\
x_3' &= -\eta\frac{x_3}{v},\\
p_1' &= -f'(x_1) + \lambda + \mu,\\
p_2' &= e^{-\delta t},\\
p_2' &= -p_2 + \frac{p_3\eta}{v},\\
v^2 &= \frac{p_3\eta x_3}{p_1}.
\end{aligned}$$

Note como la dependincia de H con respecto de v es convexa cuando $v > 0$. Este sistema de seis ecuaciones diferenciales acloplados es completado con las condiciones en los extremos y las condiciones de optimalidad,

$$\begin{aligned}
x_1(0) = k_0, \quad x_1(T) = k_T, \quad x_2(T) = u_T,\\
p_2(0) = p_3(0) = p_3(T) = 0.
\end{aligned}$$

Es imposible resolver todo el sistema de manera explicita.

No es dificil revisar que en los ejemplos anteriores las hípotesis de la suficiencia para la solución óptima estan satisfechas, asi que las soluciones encontradas son en realidad soluciones óptimas en todos los casos.

Cuando estas condiciones que aseguran optimalidad no se cumplen, entonces es posible la no existencia de soluciones óptimas. Uno de los ejemplos mas simples de esto es aquel de minimizar

$$\int_0^1 [(u(t)^2 - 1)^2 + x(t)^2]dt,$$

donde $x'(t) = u(t)$, $x(0) = 0$ y $K = \mathbb{R}$. Si tomamos

$$\begin{aligned}
u_j(t) = 1, \quad t \in \left(\frac{k}{2^j}, \frac{k+1}{2^j}\right), \quad k \text{ par },\\
u_j(t) = -1, \quad t \in \left(\frac{k+1}{2^j}, \frac{k+2}{2^j}\right), \quad k \text{ impar },
\end{aligned}$$

Es facil mostrar que $I(u_j) \searrow 0$, y por lo tanto concluimos que el valor del infimo es 0. Es sin embargo imposible encontrar un control u que tenga este valor (¿Por qué?). Note como la convexidad del integrando F con respecto al control u falla.

6.4. Otro Formato

El funcional de costo de un problema de control óptimom puede incorporar otro termino dependiendo del estado final del sistema. En general tendremos un objetivo de la forma

$$I(u) = \int_0^T F(t, x(t), u(t))dt + \phi(x(T)),$$

donde el estado final $x(T)$ es libre, la función ϕ es diferenciable y tenemos la típica ecuación de estado completada con las condiciones iniciales

$$x'(t) = f(t, x(t), u(t)), \quad x(0) = x_0.$$

Esta situación, aparentemente mas general, puede reducirse a nuestro formato típico dediante el truco

$$\begin{aligned} I(u) &= \int_0^T \left[\frac{d}{dt}\phi(x(t)) + F(t, x(t), u(t))\right] dt \\ &= \int_0^T [\nabla\psi(x(t))f(t, x(t), u(t)) + F(t, x(t), u(t))]dt. \end{aligned}$$

Por lo que es de hecho uno de nuestros ejemplos típicos con el nuevo integrando

$$\tilde{F}(t, x, u) = \nabla\phi(x)f(t, x, u) + F(t, x, u)$$

El Hamiltoniano para este nuevo problema es

$$\tilde{H}(t, x, u, p) = \nabla\psi(x)f(t, x, u) + F(t, x, u) + pf(t, x, u)$$

Note que las condiciones de óptimalidad son las mismas comparadas con aquellas para el funcional I sin el termino $\psi(x(T))$, pero con el multiplicador

$$\tilde{p}(t) = p(t) + \nabla\phi(x(t)).$$

El unico cambio es en ralidad la condicion de transversalidad que ahora es

$$\tilde{p}(T) = \nabla\phi(x(T)) \quad (p(T) = 0).$$

Y por lo tanto el cálculo de las soluciones ótimas de un problema de control de este tipo es identica al anterior, ignorando la contribución $\phi(x(T))$, la cual entra escribiendo las condiciones de transversalidad en T como se indico anteriormente.

Ejemplo 6.15 *Suponga que nos gustaria encontrar el control óptimo para minimizar el costo dado por*

$$I(u) = \frac{1}{2}x(T)^2 + \int_0^T u(t)^2 dt$$

con ecuación de estado y condiciones iniciales dadas por

$$x'(t) = -x(t) + u(t), \quad x(0) = x_0.$$

Dado que se cumplen las hipótesis para la suficiencia de las condiciones de optimalidad, podemos encontrarla aplicando el principio de Pontryagin. Como explicamos anteriormente estas condiciones son las mismas que para

$$\int_0^T u(t)^2 dt$$

pero con la condición de transversalidad correspondiete $p(T) = x(T)$*. Por lo que el Hamiltoniano es*

$$H(t, x, u, p) = u^2 + p(u - x),$$

debemos resolver el sistema

$$\begin{aligned} p' = p, \quad 2u + p = 0, \quad x' = u - x, \\ x(0) = x_0, \quad p(T) = x(T). \end{aligned}$$

despues de algunos cálculos tenemos que resolver el problema

$$x'(t) = -x(t) - \frac{1}{2}x(T)e^{t-T}, \quad x(0) = x_0.$$

Cuya solución es

$$x(t) = \frac{x_0}{5 - e^{-2T}}(5e^{-t} - e^{-t-2T}),$$

con control óptimo

$$u(t) = -\frac{2x_0 e^{-T}}{5 - e^{-2T}} e^{t-T}.$$

Bajo restricciones en el tamaño del control $u(t) \in K$*, la metodologia es similar.*

6.5. Algunos Comentarios en la Aproximación Numérica

Los problemas de control óptimo son tan importantes en ingeniería que la simulación y aproximacion numérica de estos han recibido una atención considerable, seguramente mas que los problemas variacionales. Se han analizado e implementado varias estragegias exitosas para esto (for ejemplo, metodos de dos puntos frontera, y tecnicas basadas en la condiciones de optimalidad). Esto de nuevo se encuentra mas alla del enfoque de este texto. Nuestra meta en esta sección es dar unas cuantas ideas basicas basadas directamente en discretización y optimización (no en las condiciones de óptimalidad) que nos puden ayudar en entender y apreciar el rol y las dificultades de la discretización en los problemas de control óptimo. Una buena referencia es [32], en el cual se muestra una

aproximación sistematica a los problemas de optimización incluyendo control óptimo, esta desarollado de una manera bastante completa y exhaustiva.

La aproximación numérica de problemas de control óptimo puede ser analizada como lo hicimos con los problemas variacionales, especificamente dividiendo el intervalo de tiempo en varios subintervalos, y asumiendo que los controles admisiblesm son constantes en cada uno de estos subintervalos, encontrar el mejor control factible. Cuando la partición del intervalo es lo suficientemente fina, esperamos calcular una muy buena aproximación al verdadero óptimo de nuestro problema. Esto en principio se puede hacer de esta manera, y las soluciones óptimas se pueden aproximar usando algoritmos numericos para programación no lineal. Considere la siguiente situación.

Ejemplo 6.16 *Aproximaremos numéricamente el siguiente problema de control:*

$$\textit{Minimizar} \quad I(u) = \int_0^! u(x)^2 dx$$

sujeto a

$$\begin{aligned} x''(t) = u(y), \quad u(t) \in K, \quad t \in (0,1), \\ x(0) = x'(0) = 0, \quad x(1) = 1, \quad x'(1) = 0. \end{aligned}$$

La interpretación física de este problema es clara. Nos gustaria mínimizarel gasto de combustible medido por la integral del cuadrado del control, para un objeto movil que va a viajar una distancia de 1 en linea recta y terminar con velocidad 0*. El acelerador/freno debe pertenecer a un conjunto preasignado* K*.*

Sea

$$u = (u_j), \quad j = 1, \ldots, n$$

nuestra variable independiente, donde u_j *es el valor del control en el intervalo*

$$\left(\frac{j-1}{n}, \frac{j}{n}\right).$$

Por recursión podemos resolver la ecuación de estado en cada uno de estos subintervalos. Si

$$x'((j-1)/n) = a_{j-1}, \quad x((j-1)/n) = b_{j-1},$$

son los valores funales obtenidos para x' *y* x *cuando se resuelve la ecuación de estado en el intervalo* $((j-2)/n, (j-1)/n)$*, entones debemos resolver*

$$\begin{aligned} x''(t) = u_j, \quad t \in ((j-1)/n, j/n), \\ x'((j-1)/n) = a_{j-1}, \quad x((j-1)/n) = b_{j-1}. \end{aligned}$$

y ponemos

$$a_j = x'(j/n), \quad b_j = x(j/n).$$

En esta situació simplificada todos los cálculos se pueden hacer de manera especifica, y obtenemos

$$a_j = a_{j-1} + \frac{u_j}{n}, \quad B_j = b_{j-1} + a_{j-1}\frac{1}{n} + \frac{u_j}{2n^2}.$$

Las condiciones iniciales implican $a_0 = b_0 = 0$. *Usando estas formulas recursivas apropiadamente, no es deficil confirmar que*

$$a_j = \frac{1}{n}\sum_{k=1}^{j} u_k, \quad b_j = \frac{1}{2n^2}\sum_{k=1}^{j}(2j - 2k + 1)u_k.$$

Las restricciones finales $x(1) = 1$, $x'(1) = 0$ *se traducen en*

$$\frac{1}{2n^2}\sum_{k=1}^{n}(2n - 2k + 1)u_k = 1,$$
$$\frac{1}{n}\sum_{k=1}^{n} u_k = 0.$$

Incluso podemos escribir estas dos restricciones como

$$\sum_{k=1}^{n} k u_k + n^2 = 0, \quad \sum_{k=1}^{n} u_k = 0$$

Por otro lado, el funcional de costo es simplemente

$$I(u) = \sum_{k=1}^{n} u_k^2.$$

Finalmente nos enfrentamos al problema

$$\textit{Minimizar} \quad \sum_{k=1}^{n} u_k^2$$

sujeto a

$$\sum_{k=1}^{n} k u_k + n^2 = 0, \quad \sum_{k=1}^{n} u_k = 0, \quad u_k \in K.$$

Hemos implementado dos situaciones para diferentes escogencias del conjunto K. *La primera corresponde a ninguna restricción, es decir* $k \in \mathbb{R}$. *Para la segunda hemos escogido* $K = [-1/2, 1]$. *Uno de los algoritmos numéricos del capítulo 4 puede ser usado para encontrar estas soluciones aproximadas de manera discreta. La figura 6.7 muestra la aproximación para ambos casos*

Ejemplo 6.17 *Ahora esplicamos como plantear una aproximación numérica para la solución óptima del problema 6.7.*

Dado que $T > 0$ no esta especificado (es precisamente el funcional de costo a ser minimizado), debemos incorporarlo como otra variable. Sea

$$u = (u_j), \quad j = 0, 1, \ldots, n,$$

nuestra variable independiente, donde estamos tomando $4u_0 = T$, y u_j es el valor del control en el intervalo

$$\left(u_0 \frac{j-1}{n}, u_0 \frac{j}{n}\right).$$

Recursivamente, como en el ejemplo anterior, podemos resolver la ecuación de estado en cada uno de estos subintervalos. Si

$$x'(u_0(j-1)/n) = a_{j-1}, \quad x(u_0(j-1)/n) = b_{j-1},$$

son los valores funales obtenidos para x' y x resolviendo la ecuación de estado en el intervalo $(u_0(j-2)/n, u_0(j-1)/n$, entonces tenemos que resolver

$$\begin{aligned} x''(t) = u_j, \quad t \in (u_0(j-1)/n, u_0 j/n), \\ x'(u_0(j-1)/n) = a_{j-1}, \quad x(u_0(j-1)/n) = b_{j-1}, \end{aligned}$$

y ponemos

$$a_j = x'(u_0 j/n), \quad b_j = x(u_0 j/n).$$

Como antes todos los cálculos involucrados se puden hacer de manera explicita, y asi obtenemos

$$a_j = a_{j-1} + \frac{u_0 u_j}{n}, \quad b_j = b_{j-1} + a_{j-1}\frac{u_0}{n} + \frac{u_0^2 u_j}{2n^2}.$$

Las condiciones iniciales implican $a_0 = b_0 = 0$. De nuevo no es dificil comprobar que

$$a_j = \frac{u_0}{n}\sum_{k=1}^{j} = u_k, \quad b_j = \frac{u_0^2}{2n^2}\sum_{k=1}^{j}(2j - 2k + 1)u_k.$$

Las restricciones finales se traducen en

$$\frac{u_0^2}{2n^2}\sum_{k=1}^{n}(2n - 2k + 1)u_k = \alpha, \quad \frac{u_0}{n}\sum_{k=1}^{n} u_k = 0.$$

O de manera equivalente

$$u_0^2 \sum_{k=1}^{n} k u_k + \alpha n^2 = 0, \quad \sum_{k=1}^{n} u_k = 0.$$

Por otro lado el funcional de costo se simplifica a

$$I(u) = u_0$$

Finalmente, nos enfrentamos al problema

$$\textit{Minimizar} \quad u_0$$

sujeto a

$$u_0^2 \sum_{k=1}^{n} k u_k + \alpha n^2 = 0,$$

$$\sum_{k=1}^{n} u_k = 0, \quad -l \leq u_k \leq a, \quad u_0 \geq 0.$$

Note como en los problemas de control óptimo la versión discretizada del problema fundamental de optimización requiere resolver una ecuación en diferencias. En los ejemplos examinados anteriormente, se ha hecho esto explicitamente. En muchas situaciones, incluso en situaciones simples, no podemos esperar hacer esto mismo, asi que debemos incorporar un solucionador para ecuaciones en diferencia como parde de la definición (numérica) del funcional de costo y/u de las restricciones, o de lo contrario tendriamos que ver como usar la información proveniente de las condiciones de optimalidad. Esta ultima posibilidad sera el tema de un texto mas especializado. Invitamos a nuestros lectores a formular los detalles de la aproximación numérica de otro ejemplo estandar mas elaborado, el control óptimo del oscilador armonico (Ejemplo 6.9, Ejercicio 18). Esta vez un integrador numérico (El integrador de Euler) sebe usar para aproximar la ecuación de estado sobre los subintervalos donde se debe utilizar en control constante.

6.6. Ejercicios

1. Un sistema esta gobernado por la ecuación de estado $x' + ax = u$, donde a es una constante, $x = x(t)$ es el estado, y $u = u(t)$ es el control. Si $x(0) = 0$ y $x(T) = C$, determine el control óptiimo que minimiza el costo

$$I(u) = \int_0^T \left[(C - x)^2 + u^2\right] dt.$$

   Aquí C es una constante

2. Un cierto sistema que obedece las ecuaciones de estado

$$x_1' = x_2, \quad x_2' = -x_1 + u,$$

   comienza en $x_1(0) = x_2(0) = 1$. Encuentre el control óptimo para tener el sistema en el mismo estado despues de una unidad de tiempo $x_1(1) = x_2(1) = 1$ si el costo es

$$I(u) = \frac{1}{2} \int_0^1 u^2(t) dt$$

3. Las ecuaciones
$$x_1' = x_2, \quad x_2' = -x_2 + u,$$
caracterizan el comportamiento de un sistema. Si consideramos un funcional de consto del tipo
$$I(u) = \int_0^\infty (x_1^2 + \frac{16}{3}u^2)dt,$$
encuentre el control óptimo si
$$x_1(0) = 0, \quad x_2(0) = b, \quad x_1(t), x_2(t) \to 0 \text{ mientras } t \to \infty.$$

4. Un sistema esta gobernado por la ecuación
$$x'' + x' + x = u, \quad x(0) = c_0, \quad x'(0) = c_1$$
El control esta restringido por $|u| \leq 1$. Estudie el control óptimo que lleva el sistema al reposo en un tiempo mínimo.

5. Un cohete viaja hacia arriba sobre el suejo bajo una fuerza gravitacional constante y efectos aerodinamicos despreciables. El ejector del motor actua verticalmente hacia abajo. Las ecuaciones son
$$h' = v, \quad v' = -g + \frac{c\beta}{m}, \quad m' = -\beta,$$
donde h es la altura medida con respecto al piso, v es la velocidad vertical, m es la masa total del cohete, c es una constante positiva, y β es el control representado el flujo de combustible sujeto a la restriccion $0 \leq \beta \leq \bar{\beta}$. Asumiendo que en el inicio $t = 0$, tenemos $m = m_0 + m_\beta$, donde m_β es la cantidad de combustible, $h = 0$, $v = 0$, determine el control óptimo para alcanzar la maxima altura, suponiendo que se nos permite un solo cambio en el control.

6. Una compañia decide contratar una compañia de mercadeo con el objetivo de maximizar las ventas de un cierto producto. La relación entre el nivel de ventas y la publicidad empleada, medida a traves de una función $A(t)$, esta dada por la ley
$$s' = rA\left(1 - \frac{s}{M}\right) - \lambda_s,$$
donde M es el nivel de saturación de las ventas, λ es la taza de decaimiento de las ventas si no hay publicidad, y r es un parametro positivo. Si el dinero usado en mercadeo esta limitado en cualquier momento por $0 \leq A \leq \bar{A}$ y tambien globalmente por
$$B = \int_0^T A(t)dt,$$

para un B dado, determine la estrategia optima para maximizar las ventas globales

$$S = \int_0^T s(t)dt$$

sobre un determinado periodo de tiempo.

7. Un objeto movil esta controlado por la ley

$$x'' + x' - 2x = u, \quad |u| \leq 1.$$

Si

$$x(0) = -\frac{1}{6}e^2 - \frac{1}{3}e^{-1} - \frac{1}{2}, \quad x'(0) = \frac{1}{3}e^2 - \frac{1}{3}e^{-1},$$

determine la estrategia óptima U que leva este objeto al reposo,

$$x(T) = x'(T) = 0,$$

en un tiempo mínimo. Suponga que $U(0) \leq 0$

8. Un sistema esta gobernado por la ecuación

$$x'(t) = x(t) + u(t), \quad t \in (0, 10),$$

y comienza en $x(0) = 100$. Si el costo esta dado por

$$I(u) = \frac{1}{2}x(10)^2 + \int_0^1 0 \left[3x(t)^2 + u(t)^2\right] dt,$$

determine el control óptimo.

9. Considere el circuito de la figura 6.8. La corriente inicial se anula : $i(0) = 0$. Se desea una máxima diferencia de voltaje,

$$v_0(T) = \int_0^T Ri'(t)dt$$

en el instante T. La ley del circuito es

$$i'(t) = \frac{1}{L}v_i(T) - \frac{R}{L}i(t), \quad 0 \leq v_i \leq 1.$$

Determine la estrategia óptima, y la máxima caida de potencial.

10. Dos naves A y B viajan en el espacio. En el instante inicial, estan alejadas a una distancia c_0, y B se aleja de A con una velocidad constante c_1, A quiere alcanzar a B de una manera suave. La posición de A esta gobernada por

$$x'' = u, \quad x(0) = x'(0) = 0.$$

La energía consumida por A en esta tarea es proporcional a

$$I(u) = \int_0^T u(t)^2 dt$$

Encuentre la estrategia óptima, teniendo en cuenta qeu la conexión debe ser alcanzada antes de un cierto periodo de tiempo T_1 $(T \leq T_1)$.

11. Trate de resolver el problema de control óptimo número 12 en el capítulo 1.

12. Una taza de cafe esta inicialmente a 100° Farenheit, y deseamos disminuir su temperatura en un tiempo mínimo a $0°F$ anádiendo una cantidad fija (unidad) de leche. Si $x(t)$ es la temperatura de la mezcla de cafe y leche en la taza, la ley de enfriamento de la mezcla es

$$x'(t) = -x(t) - 25u(t) - \frac{1}{4}u(t)x(t),$$

donde $u(t)$ es la razon a la cual se añade la leche, y esta restringida de tal manera que $0 \leq u(t) \leq 1$ y

$$\int_0^T u(t)dt = 1.$$

a) Argumente por quá la estrategia óptima debe tener la forma

$$u(t) = \begin{cases} 0, & 0 \leq t \leq t_0, \\ 1, & t_0 \leq T, \end{cases}$$

para un $t_0 \geq 0$ dado.

b) Teniendo en cuenta el paso anterior, encuentre la estrategia óptima

13. Una cierta plaga esta dañando un cultivo. Para eliminarla, se desarrolla un depredador y se introduce en el cultivo. Dado que el depredador es perjudicial para el cultivo, asi como infertil, se busca que ambas especies se eliminen simultaneamente tan pronto como sea posible. Si $x_1(t)$ y $x_2(t)$ designan a ambas especies, y estas comienzan desde $x_1(0) = \frac{1}{4}$, $x_2(0) = 0$, determine el control óptimo u y el tiempo mínimo si $-1 \leq u \leq 1$ y la ley de estado es

$$x_1'(t) = x_1(t) - x_2(t), \quad x_2(t) = -x_2(t) + u.$$

El contriol u representa la taza a la cual el depredador es introducido o retirado.

14. En ocasiones no se sabe si un sistema puede alcanzar un estado final deseado bajo las restricciones que tenemos. En estos casos, un problema de control óptimo como

$$\text{Minimizar} \quad \frac{1}{2}|x(T) - x_T|^2$$

sujeto a

$$x' = f(t, x, u), \quad x(0) = x_0, \quad |u| \leq M,$$

puede ayudar a determinar cuando el estado final deseado x_T puede ser alcanzado. Esto ocurre cuando en costo óptimo se anula. Decida cuando pueden haber estados inalcanzables para una ley de estado lineal con coeficientes constanes x_T.

$$f(t, x, u) = ax + bu + c, \quad a, b, c \in \mathbb{R}, \quad a, b \neq 0.$$

15. El principio de Pontryagin no siempre da soluciones óptimas. Esto es obviamente cierto cuando no hay soluciones óptimas. Trate de usar el principio del maximo de Pontryagin para el ejercicio 13 del Capítulo 1, y describa que tipo de dificultades se encuentra.

16. Examine la aproximación numérica del ejemplo 6.10 usando las ideas de la sección 6.5.

17. Estudie el problema de control óptimo de llevar el sistema gobernado por la ecuación de estado

$$x''(t) - x(t) = u(t), \quad -1 \leq u(t) \leq 1,$$

al reposo desde una posición arbitraria inicial en tiempo mínimo. Plantee el problema para la aproximación mumérica.

18. Explore la aproximación numérica del control óptimo de un oscilador armonico 6.9 usando las ideas en la sección 6.5.