

IN3701 – Modelamiento y Optimización
Auxiliar extra examen
12 de Julio, 2010**Problema 1**

- 1) Para las siguientes funciones, determine si los puntos entregados son óptimos, ya sea locales o globales. Para esto, utilice las condiciones de primer y segundo orden vistas en clase.

a) $f(x,y) = (x-2)^2 + (y-5)^3$ en el punto $(x,y)^* = (2,5)$

b) $f(x) = \frac{1}{x} + \sqrt{x}$ en el punto $x^* = 4^{1/3}$

c) $f(x) = x^{2/3} - 5$ en el punto $x^* = 0$

- 2) Para la minimización de las siguientes funciones, ¿qué método recomendaría, Newton o Gradiente?

a) $f(x,y) = x^4 + 3y^2$

b) $f(x) = x^3 + \exp(-x^2)$

c) $f(x,y) = x^2 + y^2$

- 3) Dado el siguiente problema:

$$\begin{aligned} \min f(x_1, x_2) &= (x_1 - 5)^2 + (x_2 - 2)^2 & (P) \\ \text{s.a } g_1(x_1, x_2) &= (x_1 - 1)^2 + (x_2 - 2)^2 \leq 1 \\ g_2(x_1, x_2) &= x_1 + x_2 \geq 3 \end{aligned}$$

1. Desarrolle las condiciones de Karush-Kuhn-Tucker (KKT) para el problema.
2. Revise el cumplimiento de las condiciones de KKT para los siguientes puntos:

$$(1,2); (1 - \sqrt{1/2}, 2 + \sqrt{1/2})$$

¿Qué podemos concluir para cada uno de estos puntos? Justifique.

3. Muestre las restricciones, el conjunto de soluciones factibles y la función objetivo gráficamente.
4. Determine un candidato para ser solución óptima analizando el gráfico. Verifique si este candidato cumple las condiciones de KKT. De la solución óptima y el valor de la función objetivo asociado.
5. Se agrega la siguiente restricción al problema (P):

$$g_3(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2 \leq 1$$

¿Qué podemos decir acerca de la solución óptima del problema original aplicando KKT? Justifique su respuesta.

Problema 2

- 1) Suponga que el problema de fase I del Simplex tiene óptimo = 0. ¿Es cierto siempre que la base óptima del problema auxiliar es base primal factible para el problema original? Justifique.

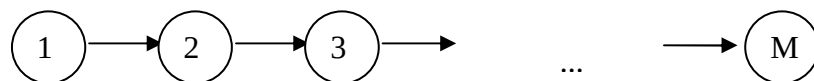
2) Explique si puede o no tener utilidad ocupar el concepto de dualidad para resolver cada uno de los problemas lineales continuos que se van definiendo en el algoritmo de ramificación y acotamiento (Branch and Bound).

3) Argumente sobre la veracidad o no de la siguiente frase: "Elegir la variable de entrada a la base como la de menor costo reducido asegura la mayor mejora de la función objetivo".

Problema 3

Una fábrica de televisores LCD, dado el éxito en ventas de sus televisores producto del mundial, ha decidido optimizar el uso de sus recursos para reducir los tiempos de producción y así generar una ventaja competitiva.

La empresa produce I modelos de televisores distintos, los cuales se deben procesar en M máquinas. Cada uno de los N modelos debe ser procesado en todas las máquinas, pasando siempre primero por la máquina 1, luego por la máquina 2, y así sucesivamente hasta terminar con la máquina M (todos los modelos pasan por las M máquinas en el mismo orden, dadas las similitudes entre ellos), tal como se muestra a continuación:



Sin embargo, algunos modelos de TV tardan más tiempo que otros en las máquinas, debido a sus características especiales. Suponga que el tiempo de proceso en la máquina m del modelo de TV i está dado por T_{im} .

1. Plantee un modelo de programación lineal entera para determinar el orden en que debe procesar los distintos modelos de televisores en las máquinas, de modo de minimizar el tiempo de tránsito total.
2. Suponga ahora que cada máquina es controlada por un operario. Suponga además que al departamento de operaciones se le ha asignado un presupuesto adicional PPTO para reducir los tiempos de producción. Usted ha decidido destinar este presupuesto para capacitar a sus trabajadores, con lo cuál reducirá en K_{im} el tiempo de proceso de un modelo de televisor i en la máquina m . La capacitación para el operario de la máquina m cuesta W_m . ¿Cómo modificaría su modelo de la parte anterior para incorporar la decisión de a qué operarios capacitar?

Solución

Problema 1

1)

a. Calculamos $\nabla f(x, y) = \begin{pmatrix} 2(x-2) \\ 3(y-5)^2 \end{pmatrix}$

Evaluando en $(x, y) = (2, 5) \Rightarrow \nabla f(2, 5) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, se cumple condición de primer orden.

Calculamos ahora $Hf(2, 5) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, es semi-definida positiva. Como **no** es definida positiva, entonces no se puede concluir nada de este punto.

b. Calculamos $\nabla f(x, y) = -\frac{1}{x^2} + \frac{1}{2\sqrt{x}}$

Evaluando en $x^* = 4^{1/3} \Rightarrow \nabla f(4^{1/3}) = -\frac{1}{4^{2/3}} + \frac{1}{2 * \sqrt{4^{1/3}}} = 0$, condición de primer orden.

Calculamos ahora $Hf(x) = \frac{2}{x^3} - \frac{1}{4 * x^{3/2}}$, evaluamos en $x^* = 4^{1/3} \Rightarrow$

$$Hf(4^{1/3}) = \frac{2}{4} - \frac{1}{4 * 2} = \frac{3}{8} > 0 \text{ es definida positiva. Con esto, se puede decir que el}$$

punto $4^{1/3}$ es mínimo local. Además, la función es convexa (deberán bosquejar un gráfico para verlo), por lo que se concluye que es mínimo **global**.

c. Calculamos $\nabla f(x) = \frac{2}{3x^{1/3}}$. No se puede evaluar en el punto $x^* = 0$, no se cumple condición de primer orden, la función no es C^1 y no es posible concluir si el punto es mínimo o no para la función con estas condiciones. Sin embargo, es fácil ver que el punto $x^* = 0$ es efectivamente el mínimo.

2) (ojo que cambia el orden de las preguntas con respecto al enunciado entregado en la clase (sección 1))

a) la función es de forma cuadrática y (estrictamente) convexa, por lo tanto newton termina en una iteración en la búsqueda del óptimo.

b) con gradiente es suficiente para encontrar el óptimo, pues la función no se caracteriza por tener comportamiento excéntrico y dado que el método es menos costoso (computacionalmente) es preferible. (poner newton no está mal tampoco, pero quitarle un poquito de puntaje, siempre que argumenten bien).

c) Tanto gradiente como newton entregan el óptimo en 1 iteración, sin embargo es más probable (e intuitivo) que digan que el método del gradiente es mejor (porque visualmente se ve que con una circunferencia se llega al óptimo en un paso).

3)

1. Primero se lleva el problema a la forma estándar

$$\begin{aligned} \min f(x_1, x_2) &= (x_1 - 5)^2 + (x_2 - 2)^2 \\ \text{s.a } g_1(x_1, x_2) &= (x_1 - 1)^2 + (x_2 - 2)^2 - 1 \leq 0 \\ g_2(x_1, x_2) &= 3 - x_1 - x_2 \leq 0 \end{aligned}$$

Así, se tiene que las condiciones de KKT son las siguientes:

$$\begin{pmatrix} 2*(x_1 - 5) \\ 2*(x_2 - 2) \end{pmatrix} + \mu_1 * \begin{pmatrix} 2*(x_1 - 1) \\ 2*(x_2 - 2) \end{pmatrix} + \mu_2 * \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Y

$$\begin{aligned} \mu_1 * ((x_1 - 1)^2 + (x_2 - 2)^2 - 1) &= 0 \\ \mu_2 * (3 - x_1 - x_2) &= 0 \end{aligned}$$

2. Punto (1,2):

$$\begin{aligned} \mu_1 * ((1-1)^2 + (2-2)^2 - 1) &= 0 \Rightarrow \mu_1 = 0 \\ \mu_2 * (3-1-2) &= 0 \Rightarrow \mu_2 \in \Re \end{aligned}$$

Entonces

$$\begin{pmatrix} 2*(1-5) \\ 2*(2-2) \end{pmatrix} + \mu_2 * \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

De aquí se obtiene:

$$\begin{aligned} \mu_2 &= -8 < 0 \\ \mu_2 &= 0 \end{aligned}$$

¡Contradicción! Entonces no cumple KKT.

Punto $(1 - \sqrt{1/2}, 2 + \sqrt{1/2})$:

$$\begin{aligned} \mu_1 * ((1 - \sqrt{1/2} - 1)^2 + (2 + \sqrt{1/2} - 2)^2 - 1) &= 0 \Rightarrow \mu_1 = \Re \\ \mu_2 * (3 - 1 + \sqrt{1/2} - 2 - \sqrt{1/2}) &= 0 \Rightarrow \mu_2 \in \Re \end{aligned}$$

entonces

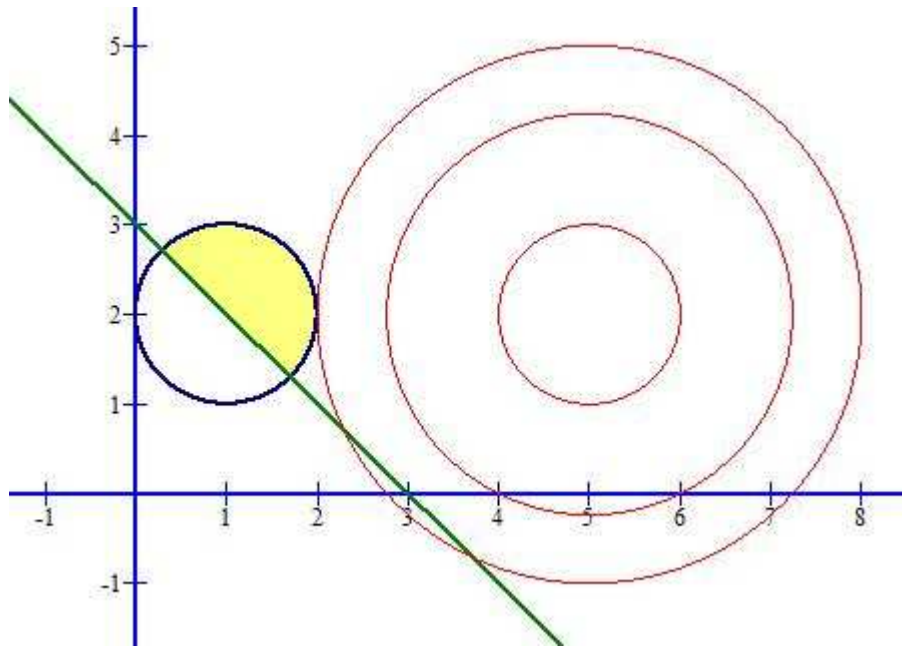
$$\begin{pmatrix} 2*(1 - \sqrt{1/2} - 5) \\ 2*(2 + \sqrt{1/2} - 2) \end{pmatrix} + \mu_1 * \begin{pmatrix} 2*(1 - \sqrt{1/2} - 1) \\ 2*(2 + \sqrt{1/2} - 2) \end{pmatrix} + \mu_2 * \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

De aquí se obtiene:

$$\begin{aligned} \mu_1 &= -(2 + \sqrt{1/2}) / \sqrt{1/2} < 0 \\ \mu_2 &= -4 < 0 \end{aligned}$$

Entonces no cumple KKT.

3.



4. Del gráfico anterior, se observa claramente que el óptimo está en el punto (2,2), por lo que proponemos dicho punto. Entonces:

$$\begin{aligned}\mu_1 * ((2-1)^2 + (2-2)^2 - 1) &= 0 &\Rightarrow \mu_1 \in \mathcal{R} \\ \mu_2 * (3-2-2) &= 0 &\Rightarrow \mu_2 = 0\end{aligned}$$

Por lo que

$$\begin{pmatrix} 2*(2-5) \\ 2*(2-2) \end{pmatrix} + \mu_1 * \begin{pmatrix} 2*(2-1) \\ 2*(2-2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

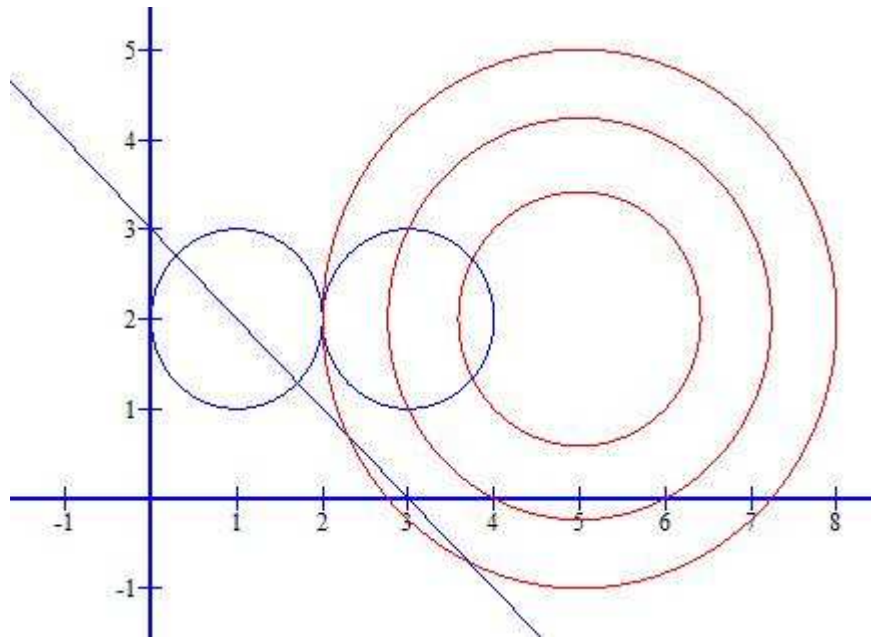
Y de aquí se obtiene:

$$\mu_1 = 3 > 0$$

¡Cumple KKT!

Como la región factible y la función objetivo f son claramente convexas, (2,2) es óptimo y $f=9$.

5. Si agregamos la restricción, el gráfico queda de la siguiente forma:



donde el punto óptimo sigue siendo el $(2,2)$, aunque ahora la región factible está formada por un único punto (el $(2,2)$, intersección de ambas restricciones circulares). Como ya probamos que éste punto cumple la condición necesaria de KKT anteriormente, y como la región factible sigue siendo convexa, se cumple KKT

(No hace falta hacer los cálculos si argumentan bien ;).

Problema 2

- 1) Cuando el valor óptimo del problema de fase I es nulo puede suceder que la base óptima del problema de fase I contiene algunas columnas asociadas a variables artificiales, pues esta es degenerada. Por lo tanto, la base óptima del problema auxiliar no es base primal factible para el problema original. Hay varias formas de poder encontrar una base primal factible en estos casos, por ejemplo, conservando las variables artificiales básicas en fase II hasta que salgan de la base o intentando eliminarlas sustituyéndolas por columnas asociadas a variables del problema original.
- 2) Al definir problemas continuos se van agregando cotas (restricciones). Al mirar el problema dual, se puede observar que este aumenta en el número de variables. Es claro que resolver problemas con aumento de variables es menos trabajoso (en cuanto a la utilización de recursos) que resolver problemas con aumento de restricciones. Luego es de utilidad resolver el problema dual.

Otra respuesta válida es: Al ir agregando las restricciones que generan los nuevos nodos del árbol de B&B, la solución del nodo padre resulta ser infactible en los nodos hijos, por lo que no es posible partir de ella para seguir iterando utilizando simplex. Sin embargo, al reconocer que la nueva restricción solo agrega una nueva variable en el problema dual, es evidente que la solución dual del nodo padre es factible en los problemas duales de los nodos hijos por lo que resulta natural utilizar el concepto de dualidad para resolver los problemas lineales que se van generando en las ramificaciones en el algoritmo B&B.

Nota: Finalmente no entra B&B en el examen así que no pesquen mucho esta pregunta :P

- 3) El criterio de entrada a la base indica que la variable no básica que entra es aquella que tiene el menor costo reducido, dentro de aquellos que son < 0 . Se ha adoptado esta convención porque trae el mayor mejoramiento local. Sin embargo al escoger esta variable de entrada se esta automáticamente determinando cual será la variable básica que saldría de la base, y puede darse el caso que esta variable aportaba "mucho" a la minimización de la función objetivo. Entonces para lograr la máxima variación habría que escoger como variable de entrada aquella que, en conjunto con la variable que saldría, impliquen la mayor variación en la función objetivo.

Problema 3:

1.

Variables: (1 pto.)

$$X_{ij} = \begin{cases} 1 & \text{si se produce modelo de TV } j \text{ después de modelo } i \\ 0 & \text{sino} \end{cases}$$

t_{im} = Tiempo de comienzo del proceso del modelo i en la máquina m

t_{FINAL} = Tiempo en que se termina de procesar el último modelo en la última máquina.

Función objetivo:

$$\min\{t_{FINAL}\}$$

Nota: Ver restricción 1

Restricciones:

1) Definición de t_{FINAL} : (0,4 ptos.)

$$t_{FINAL} \geq t_{iM} + T_{iM} \quad \forall i$$

Nota: t_{FINAL} es mayor o igual que todos los tiempos de proceso que llevan acumulados los trabajos en la máquina M. Así, el modelo i que tenga el mayor $t_{iM} + T_{iM}$ es el que le dará valor a la variable t_{FINAL} , es decir, el último trabajo que pasó por esa máquina (pues es el que tiene mayor tiempo).

2) Se debe pasar por las máquinas en orden: (0,4 ptos.)

$$t_{i,m+1} \geq t_{i,m} + T_{i,m} \quad \forall i, m \in \{1, \dots, M-1\}$$

3) Sólo se puede usar la máquina cuando está desocupada: (0,4 ptos.)

$$t_{jm} \geq t_{im} + T_{im} - (1 - X_{ij}) * M \quad \forall i, m, M \gg 0$$

Nota: Las restricciones siguientes son para definir "la ruta", i.e. asegurarse de que se produzcan todos los modelos de televisores. Es necesario agregar los nodos artificiales 0 e I+1, después se mostrará por qué. ¡Las restricciones 4,5,6 y 7 son típicas restricciones de flujo en redes!

4) De cada nodo sale un arco: (0,4 ptos.)

$$\sum_{\substack{j=0 \\ i \neq j}}^{I+1} X_{ij} = 1 \quad \forall i$$

5) A cada nodo llega un arco: (0,4 ptos.)

$$\sum_{\substack{i=0 \\ i \neq j}}^{I+1} X_{ij} = 1 \quad \forall j$$

6) Evitar subtours: (0,4 ptos.)

$$\sum_{i,j \in U} X_{ij} \leq |U| - 1 \quad \forall U \subseteq \{0,1,\dots,I,I+1\} \text{ t.q. } 2 \leq |U| \leq I-2$$

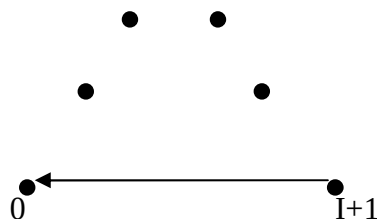
Nota: Ojo con el conjunto U, subconjunto de los nodos del 0 al I+1.

7) Condición de borde para que funcione bien la restricción 6 (existe el arco entre nodos I+1 y 0): (0,4 ptos.)

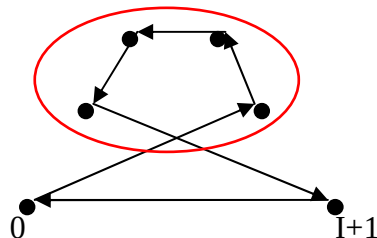
$$X_{I+1,0} = 1$$

Ejemplo: La idea es crear una ruta desde el trabajo que se realiza primero hasta el trabajo que se realiza al final, pero esta ruta no regresa al origen como en problemas de transporte o flujo en redes típicos (ej: vendedor viajero). Una vez que llegamos al último nodo nos quedamos ahí, por lo que hay que crear los nodos artificiales 0 e I+1 y obligar a que exista el arco de I+1 a 0. Con el dibujo a continuación se muestra por qué esto funciona:

Configuración inicial, donde se realizan I=4 modelos de TV distintos. Los nodos artificiales 0 e I+1=5 fueron agregados junto con el arco:



Ahora se muestra una posible configuración final factible de acuerdo a las restricciones que hemos puesto:



Se encontró una ruta entre los 4 nodos originales de la red, y NO existe un camino entre el último nodo revisado y el primero, que es lo que se quería evitar con los nodos artificiales. Los tiempos en los nodos artificiales no nos afectan en nada dado los límites de la restricción 3, para i de 1 hasta I (sin considerar los nodos artificiales).

8) Naturaleza de las variables: (0,2 ptos.)

$$X_{ij} \in \{0,1\}; t_{im} \geq 0, t_{FINAL} \geq 0 \quad \forall i, m$$

Nota de corrección: Esta explicación de las restricciones de la 4 a la 7 es sólo para que los alumnos entiendan la pauta sin tantos problemas. Obviamente no es necesario que expliquen todo esto ni que den un ejemplo como el de arriba. Si no manejan del todo bien los índices pero se dan cuenta que tienen que eliminar subtours y revisar todos los nodos (i.e. agregan las restricciones de la 4 a 7 bien, excepto por los índices) penalizar con 0,1 décimas en cada una de las que "tengan problemas de índices" (i.e. dado que es difícil de identificar este problema en el corto tiempo de una prueba, no se debe penalizar mucho).

2.

Agrego variable: (0.3 ptos.)

$$Y_m = \begin{cases} 1 & \text{si capacito al operario de la máquina } m \\ 0 & \text{sino} \end{cases}$$

Cambio las siguientes restricciones: (0,2 ptos. c/u)

- Agrego el término $-K_{im} * Y_m$ que representa el ahorro en tiempo si capacito a m
 - 1) $t_{FINAL} \geq t_{iM} + T_{iM} - K_{iM} * Y_M$
 - 2) $t_{i,m+1} \geq t_{i,m} + T_{i,m} - K_{im} * Y_m$
 - 3) $t_{jm} \geq t_{im} + T_{im} - K_{im} * Y_m - (1 - X_{ij}) * M$

El resto de las restricciones queda igual.

Agrego restricciones:

7') Naturaleza de la variable nueva: (0,2 ptos.)

$$Y_m \in \{0,1\} \quad \forall m$$

8) Respetar presupuesto: (0,4 ptos.)

$$\sum_{m=1}^M Y_m * W_m \leq PPTO$$

Función objetivo queda igual.

Dudas y/o comentarios a:
André Carboni
andre@carboni.cl