INGENIERIA INDUSTRIAL
UNIVERSIDAD DE CHILE

Profesores: Fernanda Bravo, Richard Weber, Rodrigo Wolf

Auxiliares: Andrés Abeliuk, Victor Bucarey, André Carboni, Nelson Devia

IN3701 – Modelamiento y Optimización
Control 4
19 de Noviembre, 2009

Problema 1

1) (1 pto.) 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) (1.5 pto.) Describa el método general de descenso para $f(x) \in C^1(R^n)$ partiendo de x_0. Cómo podría hacer más eficiente este método, refiérase a los parámetros que éste considera y mencione algún ejemplo basado en los métodos vistos en clase.

3) (1.0 pto.) El método de Newton se caracteriza por considerar la expansión de segundo orden de $f(x)$. Determine a partir de ésta la expresión para la dirección de descenso que utiliza este método de optimización.

4) (0.5 pto.) Para la minimización de las siguientes funciones, ¿qué método recomendaría, Newton o Gradiente?
   a) $f(x,y) = x^4 + 3y^2$
   b) $f(x) = x^3 + \exp(-x^2)$

5) (2 ptos.) Para los siguientes problemas de optimización (P) y (P')

$$(P) \quad Min F(x,y): \frac{x^2}{16} + \frac{(y+3)^2}{9}$$
$$g_1(x,y) = x^2 - y \leq 0$$
$$g_2(x,y) = x^2 + (y-2)^2 \leq 4$$

$$(P') \quad Min F(x,y): (x-1)^2 + (y+1)^2$$
$$g_1(x,y) = 15x^3 - 5y \leq 0$$

   a) Desarrolle las condiciones de KKT para ambos problema. Evalúe las condiciones en el punto (0,0) y analice su cumplimiento. Analice optimalidad. Justifique.
   b) Grafique los problemas (P) y (P') e identifique el óptimo gráficamente.

Problema 2

La conocida tienda por departamentos ZIPPY & CO. desea determinar su plan de abastecimiento y distribución para los próximos T días. La empresa cuenta con un centro de distribución único, y con I tiendas a lo largo de todo Chile, las que ofrecen N productos diferentes.

El centro de distribución recibe productos de sus proveedores, los que son enviados inmediatamente a la zona de despachos, donde son procesados y despachados a las respectivas tiendas, o bien son enviados a bodega. Esta bodega se encuentra ubicada en el mismo centro de distribución y cuenta con un espacio de $H[m^3]$. Considere que cada unidad de producto n tiene un volumen $F_n[m^3/unidad]$. El centro de distribución tiene, para cada período t, una capacidad de recepción de $MaxR_t$ y una capacidad de despacho de $MaxD_t$ unidades de cualquier producto, esto debido al espacio y mano de obra acotados. Considere que los costos de procesamiento en el centro de distribución el día t son de α_t [$/unidad] y el costo diario de almacenamiento en bodega es de $M[\$/m^3]$. Al comienzo del horizonte, el centro de distribución cuenta con R_n unidades de producto n en bodega.

Debido a las limitantes de espacio y mano de obra, se tiene la opción de arrendar por día alguna de las K posibles bodegas externas. Estas bodegas no tienen la capacidad de almacenar en inventario, pero sirven para procesar productos que la empresa recibe de su proveedor y despacharlos a las tiendas durante el mismo día, con el objetivo de disminuir la carga de trabajo en el centro de distribución. Cada bodega externa k tiene capacidad para procesar hasta MaxP unidades al día de cualquier producto a un costo unitario β_k [$/unidad]. El costo de arriendo de la bodega k es A_k[$] por día.

El costo de transporte desde el centro de distribución a la tienda i es de $\mu_i[\$/m^3]$ y desde la bodega externa k a la tienda i es de $\lambda_{ki}[\$/m^3]$. Los envíos desde el centro de distribución o cualquier bodega externa hasta la tienda i demoran δ_i días en llegar. Además, cada tienda i ha acordado el envío de D_{nit} unidades del producto n para cada día $t>\delta_i$ (es decir, los primeros δ_i días, la demanda de productos por parte de la tienda i es cero). Considere además que el centro de distribución puede dejar demanda insatisfecha por motivos operacionales. El costo de dejar demanda insatisfecha del producto n en la tienda i es de θ_{in} [$/unidad].

Suponga que cada producto n es adquirido a un único proveedor (es decir, hay N proveedores, uno para cada producto) a un costo unitario C_n[$/unidad], siendo los proveedores los responsable de transportar los productos al centro de distribución o bodegas externas arrendadas en el período. Los proveedores han informado que para el período t, puede vender a la empresa hasta B_{nt} unidades del producto n.

Plantee un modelo de programación lineal mixto, que minimice los costos de abastecimiento y distribución de ZIPPY & CO. en el horizonte de T días.

Pauta:

Pregunta 1

1)

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

Evaluando en (x,y)=(2,5) => $\nabla f(2,5)=\begin{pmatrix}0\\0\end{pmatrix}$, 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 $\nabla f(x,y)=-\dfrac{1}{x^2}+\dfrac{1}{2\sqrt{x}}$

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

Calculamos ahora $Hf(x)=\dfrac{2}{x^3}-\dfrac{1}{4*x^{3/2}}$, evaluamos en x*= $4^{1/3}$ => $Hf(4^{1/3})=\dfrac{2}{4}-\dfrac{1}{4*2}=\dfrac{3}{8}>0$ 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 $\nabla f(x)=\dfrac{2}{3x^{1/3}}$. 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)

Require: $x^o \in \mathbb{R}^n$, $f \in \mathcal{C}^1(\mathbb{R}^n)$, $\varepsilon > 0$.

$k \leftarrow 0$, $\delta^k \leftarrow \|\nabla f(x^k)\|$.

while $\delta^k > \varepsilon$ **do**

  Buscar $x^{k+1} : f(x^k) > f(x^{k+1})$.

  $\delta^{k+1} \leftarrow \|\nabla f(x^{k+1})\|$.

  $k \leftarrow k+1$.

return $x^* = x^k$.

Se busca un punto estacionario de una función escalar *f*. Se generan iterativamente una sucesión de puntos $\{x^k\}$; *k*=0, 1, 2... tal que la sucesión correspondiente $\{f(x^k)\}$; *k*=0, 1, 2... sea monótona decreciente. Una sucesión que verifica esta propiedad se dice que es una *sucesión de descenso* para *f*.

Este método general de descenso podría hacerse más eficiente en la forma que se determina el siguiente punto X^{k+1} en cada iteración. En clase se vio que éste se determina según $X^{k+1}=x^k+cd$, luego escogiendo el paso c y la dirección de descenso d de manera “inteligente”, podrá aumentarse la eficiencia del método. Por ejemplo, en gradiente la dirección de descenso corresponde a la dirección de máximo descenso de f, $-\nabla f(x)$ y el paso λ_k como el paso que minimiza el valor de $f(x^k+\lambda_k d)$. Para newton, la dirección de descenso es $-Hf(x)^{-1}\nabla f(x)$, y el paso es unitario. En el algoritmo de búsqueda binaria, la dirección es dada (pero debe ser de descenso) y el paso se escoge de manera de “encerrar” al óptimo en un intervalo y en cada iteración va “achicándolo” según la exactitud que se le exija.

3) Por Taylor de segundo orden, $f(x+d)=f(x)+\nabla f(x)^t d+\frac{1}{2}d^t\nabla^2 f(x+\lambda d)d$ para algún $\lambda\in(0,1)$. Por continuidad, y para d pequeños, aproximamos $\nabla^2 f(x+\lambda d)$ a $\nabla^2 f(x)$.

Asumiendo $\nabla^2 f(x)$ >0, minimizamos $f(x+d)=f(x)+\nabla f(x)^t d+\frac{1}{2}d^t\nabla^2 f(x)d$.

- El óptimo es $d=-(\nabla^2 f(x)^{-1})\nabla f(x).$
- Notemos que $d^t\nabla f(x)=-\nabla f(x)^t(\nabla^2 f(x))^{-t}\nabla f(x)$
- i.e. si $\nabla^2 f(x)>0$, d es dirección de descenso.

4)
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).

5) *Problema (P):*

Primero se lleva el problema a la forma estándar

$$
\begin{aligned}
(P)\qquad & MinF(x,y):\frac{x^2}{16}+\frac{(y+3)^2}{9}\\
& g_1(x,y)=x^2-y\le 0\\
& g_2(x,y)=x^2+(y-2)^2-4\le 0
\end{aligned}
$$

Ahora calculamos los gradientes de f y g_i

$$\nabla f=\begin{pmatrix}2x/16\\2(y+3)/9\end{pmatrix}\ ;\ \nabla g_1=\begin{pmatrix}2x\\-1\end{pmatrix}\ ;\ \nabla g_2=\begin{pmatrix}2x\\2(y-2)\end{pmatrix}$$

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

$$\begin{pmatrix}2x/16\\2(y+3)/9\end{pmatrix}+\mu_1*\begin{pmatrix}2x\\-1\end{pmatrix}+\mu_2*\begin{pmatrix}2x\\2(y-2)\end{pmatrix}\quad=0$$

y

$$\mu_1 * \ (x^2 - y) = 0$$
$$\mu_2 * (x^2 + (y-2)^2 - 4) = 0$$

Evaluemos el punto (0,0), nos queda:

$$\mu_1 * \ (0-0) = 0 \qquad \Rightarrow \mu_1 \in \Re$$
$$\mu_2 * (0 + (0-2)^2 - 4) = 0 \qquad \Rightarrow \mu_2 \in \Re$$

Entonces,

$$\begin{pmatrix} 0 \\ 6/9 \end{pmatrix} + \mu_1 * \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \mu_2 * \begin{pmatrix} 0 \\ -4 \end{pmatrix} = 0$$

A primera vista no podemos despejar los valores de los multiplicadores, sin embargo debemos notar que $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$ y $\begin{pmatrix} 0 \\ -4 \end{pmatrix}$ son linealmente dependientes, ***¡*el punto no es regular*!* Luego, no podemos concluir nada de este punto con KKT.** Sin embargo, gráficamente se verá en la parte b) que éste sí es el óptimo del problema.

<u>*Problema (P'):*</u>

Notamos que el problema ya está en forma estándar:

$$(P') \qquad MinF(x,y) : (x-1)^2 + (y+1)^2$$
$$g_1(x,y) = 15x^3 - 5y \leq 0$$

Ahora calculamos los gradientes de f y g_i

$$\nabla f = \begin{pmatrix} 2(x-1) \\ 2(y+1) \end{pmatrix} \quad ; \quad \nabla g_1 = \begin{pmatrix} 45x^2 \\ -5 \end{pmatrix}$$

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

$$\begin{pmatrix} 2(x-1) \\ 2(y+1) \end{pmatrix} + \mu_1 * \begin{pmatrix} 45x^2 \\ -5 \end{pmatrix} = 0$$

y

$$\mu_1 * \ (15x^3 - 5y) = 0$$

Evaluemos el punto (0,0), nos queda:

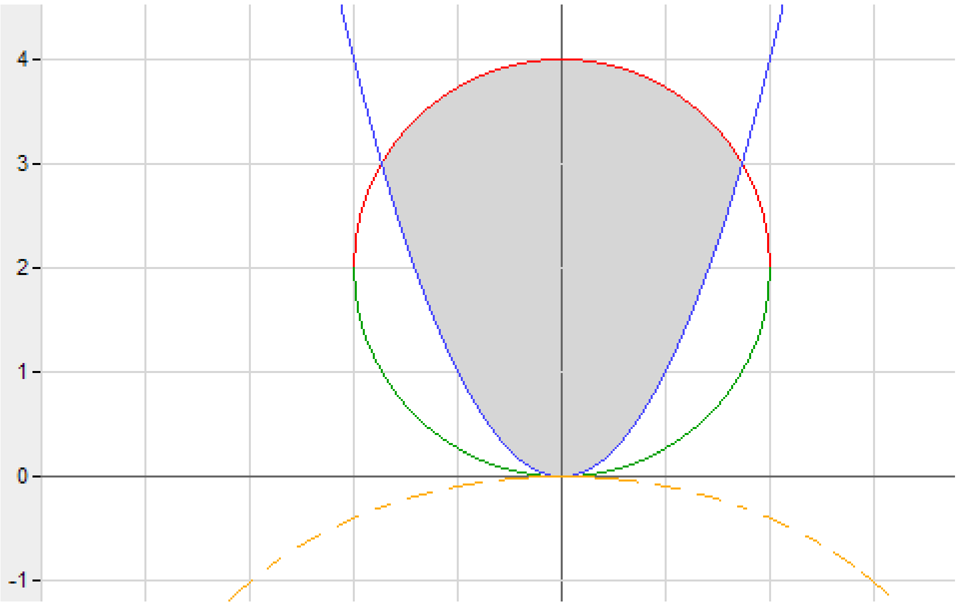
$$\mu_1 * \ (0-0) = 0 \qquad \Rightarrow \mu_1 \in \Re$$

Entonces,

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

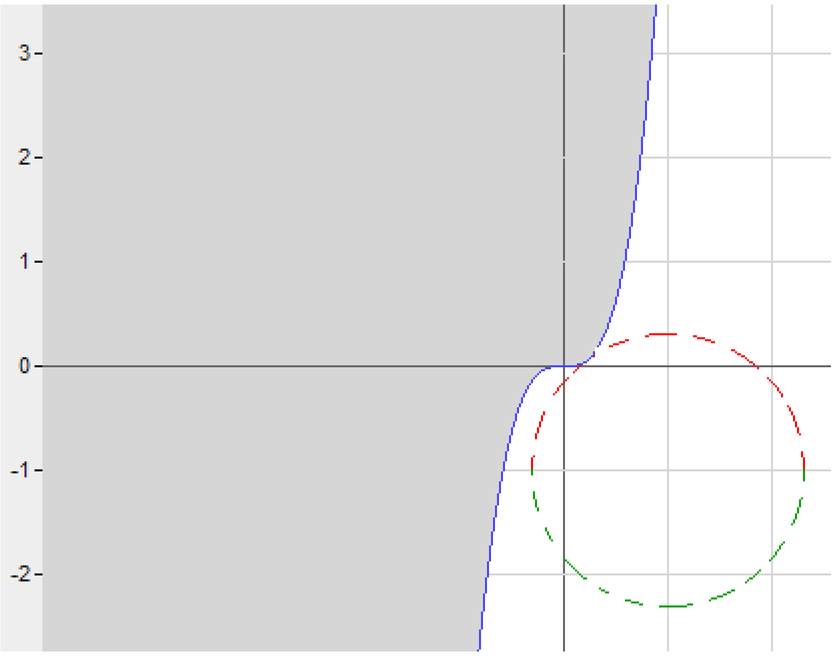
De aquí -2=0, ¡contradicción! Esto se debe a que la región no es convexa, luego el punto no cumple KKT. Es lógico que no lo cumpla, ya que gráficamente se verá en b) que el punto (0,0) no es el óptimo.

b) Grafico problema (P):



El óptimo está en el (0,0).

Gráfico problema (P’):



Aquí no es necesario que expliciten cuál es el óptimo (pues las ecuaciones son difíciles), pero sí deben darse cuenta de que el óptimo no es (0,0) como se daba a entender con la parte A) cuando se pidió KKT.

Pregunta 2

a)
Variables de decisión: (1 pto.)

X_{nt} = Unidades compradas de producto n en período t.
$\overline{X}_{nkt}$ = Unidades compradas del producto n que son enviadas a la bodega externa k en período t.
Y_{nit} = Unidades del producto n enviadas a la tienda i en período t.
$\overline{Y}_{nkit}$ = Unidades del producto n enviadas desde bodega externa k a la tienda i en período t.
I_{nt} = Unidades del producto n en inventario en el Centro de distr. en t.
S_{nit} = Unidades de demanda insatisfecha del producto n en tienda i en t.

$$Z_{kt} = \begin{cases} 1 & \text{Si se arrienda bodega externa k en t} \\ 0 & \sim \end{cases}$$

Restricciones:

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

$$X_{nt}, \overline{X}_{nkt}, Y_{nit}, \overline{Y}_{nkit}, I_{nt}, S_{nit} \in \mathbb{N}$$
$$Z_{kt} \in \{0,1\}$$

2. Satisfacer disponibilidad de productos: (0,5 ptos.)

$$X_{nt} + \sum_{k} \overline{X}_{nkt} \leq B_{nt} \qquad \forall n,t$$

3. Satisfacer la demanda de cada tienda: (0,5 ptos.)

$$Y_{ni,t-\delta i} + \sum_{k} \overline{Y}_{nki,t-\delta i} \geq D_{nit} - S_{nit} \qquad \forall n,i \quad \forall t \text{ tal que } \delta_i < t \leq T$$

4. Capacidad de almacenamiento del centro de distribución: (0,5 ptos.)

$$\sum_{n} F_n I_{nt} \leq H \qquad \forall t$$

5. Capacidad de recepción del centro de distribución: (0,4 ptos.)

$$\sum_{n} X_{nt} \leq MaxR_t \qquad \forall t$$

6. Capacidad de despacho del centro de distribución: (0,4 ptos.)

$$\sum_{n,t} Y_{nit} \leq MaxD_t \qquad \forall t$$

7. Ecuación de inventario del centro de distribución: (0,5 ptos.)

$$I_{n,t-1} + X_{nt} - S_{ni,t-1} = I_{nt} + \sum_i Y_{nit} - S_{nit} \qquad \forall n \ \ \forall t > 1$$

no olvidarse de la condición de borde:

$$R_n + X_{nt} = I_{n1} + \sum_i Y_{ni1} - S_{ni1} \qquad \forall n$$

(podrían definir también I_0=R_n y S_0=0 en vez de la anterior, y restricción 7 queda "para todo t").

8. Continuidad de flujo de las bodegas externas: (0,5 ptos.)

$$\sum_n \overline{X}_{nkt} \leq MaxP \cdot Z_{kt} \qquad \forall k, t$$

9. Capacidad de procesamiento de cada bodega externa: (0,5 ptos.)

$$\overline{X}_{nkt} = \sum_i \overline{Y}_{nkit} \qquad \forall n \ \ \forall t$$

Función objetivo: (1 pto.)

$$\min Z = \left(\sum_{n,t} C_n \cdot X_{nt} + \sum_{n,k,t} C_n \cdot \overline{X}_{nkt} \right) + \left(\sum_{n,i,t} \mu_i \cdot F_n \cdot Y_{nit} + \sum_{n,i,k,t} \lambda_{ki} \cdot F_n \cdot \overline{Y}_{nkit} \right) + \left(\sum_{n,i,t} \alpha_t \cdot Y_{nit} + \sum_{n,i,k,t} \beta_k \cdot \overline{Y}_{nkit} \right)$$

$$+ \sum_{n,t} M_t \cdot F_n \cdot I_{nt} + \sum_{k,t} A_k \cdot Z_{ktt} + \sum_{n,i,t} \theta_{nt} \cdot S_{nit}$$

Nota de corrección: En algunas restricciones pudieron haber puesto $\forall t < T - \delta_i$*, esto porque como el envío tarda* δ_i *días en llegar, no hace falta decidir el envío de productos después de* $T - \delta_i$ *días. Si lo pusieron en las restricciones correctas (las relacionadas con envío de productos a las tiendas), está bien.*

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