



Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Industrial
IN34A – Optimización

Profesores: Leonardo López, Ximena Schultz,
Richard Weber
Auxiliares: Fernanda Bravo, André Carboni,
Nelson Devia, Ignacio Escobar,
Rodrigo Wolf

Auxiliar N°2 **13 de Agosto, 2008**

Problema 1

Responda las siguientes preguntas sobre métodos de descenso:

1. Describa una iteración del Método del Gradiente y dé una justificación gráfica y/o conceptual de su funcionamiento.
2. Describa una iteración del Método de Newton y dé una justificación gráfica y/o conceptual de su funcionamiento.

Problema 2:

Considere el siguiente problema de optimización:

$$\text{Min } f(x_1, x_2) = x_1^2 + x_2^2$$

Resolver utilizando el método del Gradiente y, luego, el de Newton. Tome como punto de partida el (2,1).

Problema 3:

Considere el siguiente problema de optimización:

$$\text{Min } f(x) = x^4 - 24x^2$$

Resolver utilizando el método de Newton. Tome como punto de partida el $x = 1$. ¿Qué puede observarse?

Problema 4:

Considere el siguiente problema de optimización:

$$\text{Min } f(x_1, x_2) = x_1^2 + \frac{1}{2}x_2^2$$

Resolver utilizando el Método de Newton y del Gradiente. Tome como punto de partida el (2,4). ¿Qué puede observarse de ambos métodos?

Problema 5:

Nombre las principales ventajas y desventajas de ambos métodos.

Problema 6:

El nuevo programa de postgrado MPA de una prestigiosa universidad ha recibido n postulaciones excelentes para su primera generación y por capacidad no puede aceptar todas. Tiene el problema de seleccionar las hasta m postulaciones más adecuadas para el programa ($m < n$).

Como es de carácter público no puede solamente seleccionar para maximizar el beneficio económico del programa. Cada generación tiene que cumplir además con los siguientes criterios.

Por lo menos 40% de los aceptados tienen que pertenecer a cada sexo. El parámetro S_i indica el sexo del postulante i ($i=1, \dots, n$). $S_i = 1$ significa femenino y $S_i = 0$ significa masculino.

El programa busca tener impacto en regiones por lo que por lo menos 50% de los aceptados de cada generación deben venir de regiones. El parámetro R_i indica la proveniencia del postulante i ($i=1, \dots, n$). $R_i = 1$ significa regiones y $R_i = 0$ significa RM.

Es importante lograr alta exigencia académica por lo que el programa quiere aceptar alumnos que en promedio tengan un puntaje de por lo menos $P_{\min}$. P_i es el puntaje del postulante i ($i=1, \dots, n$).

El ingreso generado por cada generación debe superar en por lo menos 20% los costos asociados C Unidades Monetarias (UM).

En un principio cada alumno aceptado tiene que pagar un arancel de A UM.

El programa ofrece hasta B becas para alumnos excelentes ($B < n$). Los criterios para otorgar estas becas son los siguientes.

Un postulante i con un puntaje P_i mayor que 750 paga sólo 50% del arancel si viene de la RM y paga sólo 25% del arancel si viene de regiones. .

El objetivo es maximizar el ingreso generado por cada generación aceptada.

1. Plantee un modelo de Programación Lineal para resolver este problema.
2. ¿Son todas las restricciones mencionadas necesarias? Comente.
3. ¿Se está favoreciendo a alguien con la formulación actual?

Solución:

Problema 1:

1. Iteración k del Método del Gradiente:

- Calcular el gradiente de la función objetivo y verificar si $\nabla f(x^k) = 0$.
- Si no, obtener el siguiente punto como se indica a continuación:

$$x^{k+1} = x^k - \lambda_k \nabla f(x^k)$$

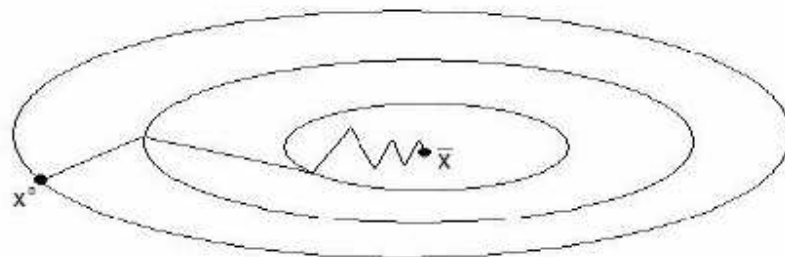
- Donde λ_k es la solución óptima del problema:

$$\min_{\lambda} h_k(\lambda) = f(x^k + \lambda d^k)$$

$$\text{s.a. } \lambda \geq 0$$

- Si es, la entonces se acaba iteración y el punto x^k es el óptimo.

Este método se basa en el hecho que la dirección de máximo ascenso de una función esta dado por su gradiente y por tanto la de mayor descenso viene dada por menos (-) el gradiente. Observar que nos acercamos zig-zageando al óptimo pues los gradientes de iteraciones sucesivas son ortogonales.



2. Iteración k del Método de Newton:

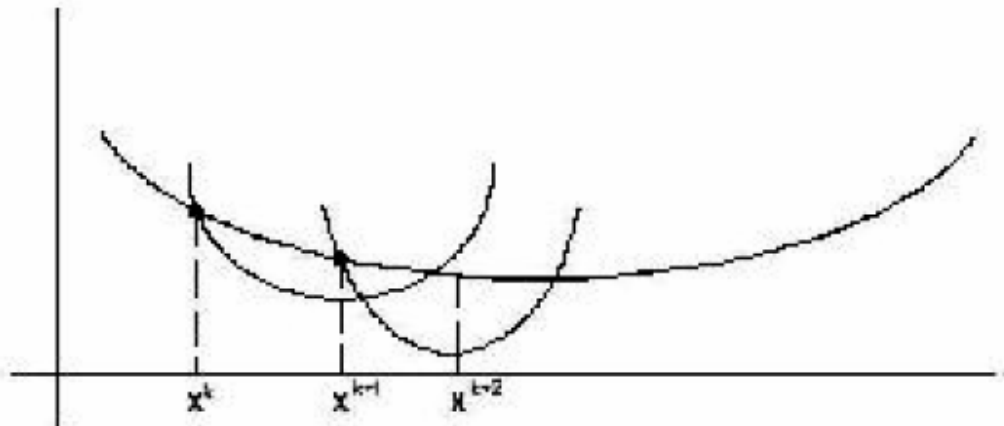
- Calcular el gradiente de la función objetivo y verificar si $\nabla f(x^k) = 0$.
- Si no, calcular el inverso del hessiano de la función objetivo para obtener el siguiente punto como se indica a continuación:

$$x^{k+1} = x^k - [H f(x^k)]^{-1} \nabla f(x^k)$$

- Si es, entonces se acaba la iteración y el punto x^k es el óptimo.

El Método de Newton se basa en aproximar la función f por una función cuadrática (en cada punto x^k se aproxima por la expansión de Taylor de orden 2) y esta aproximación se minimiza exactamente, generando un nuevo punto x^{k+1} .

Se detiene la aproximación cuando se llega a un punto estacionario de la función aproximada lo que es condición necesaria y suficiente.



Problema 2:

- Método del gradiente: $\vec{x}_{k+1} = \vec{x}_k - \lambda_k \nabla f(\vec{x}_k)$

Claramente, antes de comenzar a iterar se requiere calcular $\nabla f(\vec{x})$:

$$\nabla f(\vec{x}) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

Ahora, comenzamos a iterar:

- Iteración 1:

- Vemos si estamos en el óptimo: $\nabla f(2, 1) = (4, 2) \neq \vec{0} \Rightarrow$ Seguimos.
- Calculamos nuevo punto $\vec{x}_1$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \lambda_0 \begin{pmatrix} 4 \\ 2 \end{pmatrix}$$

Para calcular λ_0 minimizamos $h(\lambda)$:

$$\begin{aligned} h(\lambda) &= f(\vec{x}_0 - \lambda \nabla f(\vec{x}_0)) \\ &= f\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \lambda \begin{pmatrix} 4 \\ 2 \end{pmatrix} \right\} \\ &= (2 - 4\lambda)^2 + (1 - 2\lambda)^2 \end{aligned}$$

$$\frac{dh(\lambda)}{d\lambda} = 0 \Rightarrow \lambda_0 = \frac{1}{2}$$

(Si quieren pueden verificar que el punto crítico es óptimo comprobando que $d^2h/d\lambda^2 > 0$)

Entonces:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \vec{x}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

• Iteración 2:

◦ Vemos si estamos en el óptimo: $\nabla f(0,0) = (0,0) = \vec{0} \Rightarrow \text{Fin.}$

$\therefore \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ es el mínimo de $x_1^2 + x_2^2$

■ Método de Newton: $\vec{x}_{k+1} = \vec{x}_k - \{H_f(\vec{x}_k)\}^{-1} \nabla f(\vec{x}_k)$

Claramente, antes de comenzar a iterar se requiere calcular $\nabla f(\vec{x})$ y $\{H_f(\vec{x})\}^{-1}$:

$$\nabla f(\vec{x}) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

$$H_f(\vec{x}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \Rightarrow \{H_f(\vec{x})\}^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

Ahora, comenzamos a iterar:

• Iteración 1:

◦ Vemos si estamos en el óptimo: $\nabla f(2,1) = (4,2) \neq \vec{0} \Rightarrow \text{Seguimos.}$

◦ Calculamos nuevo punto $\vec{x}_1$:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \vec{x}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

• Iteración 2:

◦ Vemos si estamos en el óptimo: $\nabla f(0,0) = (0,0) = \vec{0} \Rightarrow \text{Fin.}$

$\therefore$ También se concluye que $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ es el mínimo de $x_1^2 + x_2^2$

Problema 3:

Calculamos el gradiente y el hessiano:

$$\nabla f(x) = 4x^3 - 48x$$

$$Hf(x) = 12x^2 - 48 \Rightarrow Hf^{-1}(x) = \frac{1}{(12x^2 - 48)}$$

- Iteración 0:
 - Vemos si estamos en el óptimo: $\nabla f(x=1) = 4 - 48 = -44 \neq 0$, entonces seguimos.
 - Calculamos el segundo punto: $x_1 = 1 - \left(\frac{-1}{36}\right)(-44) = -\frac{2}{9}$.
- Iteración 1:
 - Vemos si estamos en el óptimo: $\nabla f(x = -\frac{2}{9}) = \frac{7744}{729} \neq 0$, entonces seguimos.
 - El segundo punto: $x_2 = -\frac{2}{9} - \left(\frac{1}{12\left(\frac{2}{9}\right)^2 - 48}\right)\left(\frac{7744}{729}\right) = \frac{1}{540}$.
- Iteración 2:
 - Vemos si estamos en el óptimo: $\nabla f(x = \frac{1}{540}) = \frac{-4}{45} \neq 0$, entonces seguimos.
 - El segundo punto: $x_3 = \frac{1}{540} - \left(\frac{1}{12\left(\frac{1}{540}\right)^2 - 48}\right)\left(\frac{-4}{45}\right) = 0$.
- Iteración 3:
 - Vemos si estamos en el óptimo: $\nabla f(x=0) = 0$, por lo tanto se acaba la iteración.
 - El hessiano esta dado por: $Hf(x) = 12(0)^2 - 48 = -48$

Iteración	Punto (x)	Gradiente (f'(x))	Hessiano (f''(x))	Funcion objetivo
0	1	-44	-36	-23
1	- 2/9	7744/729	-1280/27	-1042/881
2	1/540	- 4/45	-48	0
3	0	0	-48	0

Como se puede observar el método hace crecer a la función objetivo, convergiendo a un máximo, esto es porque el punto x_0 no fue bien elegido dado que el hessiano de f es definido negativa en el intervalo $(-2,2)$.

Problema 4:

Método de Newton:

Iteración 1:

Primero calculamos el gradiente: $\nabla f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ x_2 \end{pmatrix}$

Luego el Hessiano: $Hf \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \Rightarrow Hf^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}$

El punto de partida es: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$

Evaluamos el gradiente en el punto: $\nabla f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, lo que implica que no

estamos en el óptimo.

Calculamos un segundo punto a partir del método:

$$x^1 = \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 4 \end{pmatrix}$$

El siguiente punto es:

$$x^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Iteración 2:

Evaluamos el gradiente en el punto: $\nabla f(x^1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \text{FIN.}$

El óptimo es: $x^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Método del Gradiente:

Iteración 1:

Primero calculamos el gradiente: $\nabla f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ x_2 \end{pmatrix}$

El punto de partida es: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$

Evaluamos el gradiente en el punto: $\nabla f \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, lo que implica que no

estamos en el óptimo.

Calculamos un segundo punto a partir del método:

$$x^1 = \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \lambda_0 \nabla f \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

$$\text{con } \lambda_k \rightarrow \min_{\lambda} f(x^k - \lambda \nabla f(x^k))$$

$$f(x^0 - \lambda_0 \nabla f(x^0)) = f \begin{pmatrix} 2 - \lambda_0 4 \\ 4 - \lambda_0 4 \end{pmatrix} = (2 - \lambda_0 4)^2 + \frac{(4 - \lambda_0 4)^2}{2}$$

$$\lambda_0 \rightarrow \min (2 - \lambda_0 4)^2 + \frac{1}{2} (4 - \lambda_0 4)^2$$

$$\frac{\partial ((2 - \lambda_0 4)^2 + \frac{1}{2} (4 - \lambda_0 4)^2)}{\partial \lambda_0} = 0 \Rightarrow \lambda_0 = \frac{2}{3}$$

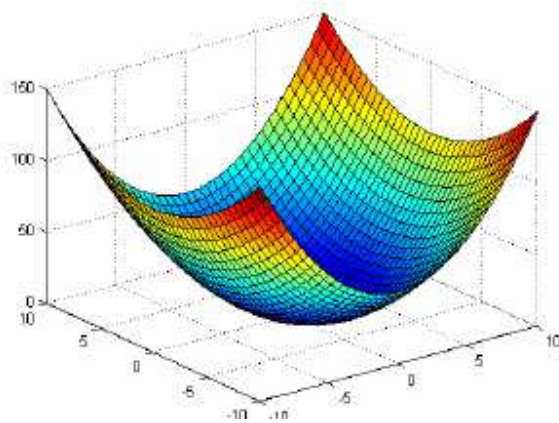
El siguiente punto es:

$$x^1 = \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \frac{2}{3} \begin{pmatrix} 4 \\ 4 \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} \\ \frac{4}{3} \end{pmatrix}$$

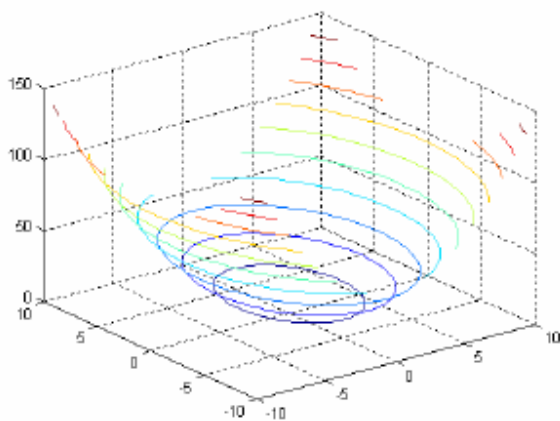
Iteración 2:

etc...

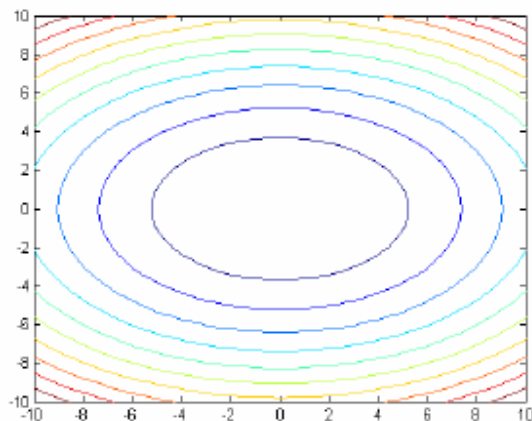
Vemos que la convergencia del método de Newton en este caso es mejor que la del método del Gradiente. Esto es, con el método de Newton se converge más rápido al punto. Esto se debe a que las superficies de nivel son excéntricas lo que se traduce en un mayor número de iteraciones en el método del gradiente. Los gráficos siguientes fueron generados con Matlab e ilustran la forma de elipsoide de las curvas de nivel de esta función.



```
[x,y]=meshgrid(-10:.5:10);
z=x.^2+(1/2)*y.^2;
plot3(x,y,z)
mesh(x,y,z)
surf(x,y,z)
```



```
[x,y]=meshgrid(-10:.5:10);
z=x.^2+(1/2)*y.^2;
contour3(x,y,z,10)
```



```
[x,y]=meshgrid(-10:.5:10);
z=x.^2+(1/2)*y.^2;
contour(y,x,z,10)
```

Problema 5:

En general el método de Newton converge más rápido que el del gradiente, este último es muy bueno cuando las superficies de nivel son hiperesferas (circunferencias en dos o más dimensiones), en tal caso el gradiente encuentra el mínimo en una iteración, pero si las superficies de nivel son excéntricas la convergencia puede ser muy lenta.

Otra característica del método del gradiente es que siempre irá orientándose hacia el punto buscado, lo que no necesariamente ocurre con el método de Newton. Otro problema de este es que hay que calcular el Hessiano e invertirlo, lo que puede ser complicado.

Una buena técnica sería partir con el método del gradiente, y seguir con el de Newton.

Problema 6:

1. Variables de decisión:

$$X_i = \begin{cases} 1 & \text{Si se acepta al alumno } i \\ 0 & \sim \end{cases}$$

$$Y_i = \begin{cases} 1 & \text{Si alumno } i \text{ recibe beca} \\ 0 & \sim \end{cases}$$

Nota: si el alumno pone una variable 1 si tienen mas de 750 pts y cero sino, sólo tendrá el puntaje en la restricción de asignar las becas (h) si relaciona dicha variable con el puntaje mediante una restricción y dicha restricción está bien planteada.

Restricciones:

a. Relación entre variables:

$$Y_i \leq X_i$$

b. Se aceptan a lo más m postulantes:

$$\sum_{i=1}^n X_i \leq m$$

c. Al menos 40% de los aceptados deben ser de cada sexo:

$$0,4 * \sum_{i=1}^n X_i \leq \sum_{i=1}^n X_i * S_i \leq 0,6 * \sum_{i=1}^n X_i$$

d. 50% de los aceptados deben venir de regiones:

$$0,5 * \sum_{i=1}^n X_i \leq \sum_{i=1}^n X_i * R_i$$

e. Puntaje promedio de los alumnos debe ser al menos Pmin:

$$P_{\min} * \sum_{i=1}^n X_i \leq \sum_{i=1}^n X_i * P_i$$

f. Ingreso generado debe superar en al menos 20% los costos:

$$\sum_{i=1}^n X_i * A - \sum_{i=1}^n Y_i * R_i * 0,75 * A - \sum_{i=1}^n Y_i * (1 - R_i) * 0,5 * A \geq 1,2 * C$$

g. Máximo B becas:

$$\sum_{i=1}^n Y_i \leq B$$

h. Criterio para otorgar becas:

$$(X_i - Y_i) * P_i \leq 750$$

Explicación: Si el puntaje de un postulante i es P_i mayor que 750 (por ejemplo 800), esto obliga a que $(X_i - Y_i) = 0$ para que se cumpla la restricción, con lo cual $X_i = Y_i = 0$ ó $X_i = Y_i = 1$ (El caso $X_i = 0$ e $Y_i = 1$ también es factible, ya que $-800 < 750$ cumple la restricción, pero éste se descarta por la restricción "a"), con lo que obligatoriamente debemos darle beca al postulante i si lo aceptamos.

En caso de que un P_i sea menor que 750 (por ejemplo 700), el término $(X_i - Y_i)$ puede valer tanto 0 como 1 para que se cumpla la restricción, con lo que se podría llegar a pensar que le estamos dando beca a los alumnos con puntaje inferior a 750 (caso $X_i = Y_i = 1$). Sin embargo hay que notar que al resolver el PPL, al estar maximizando el ingreso, el programa que lo resuelve tratará de asignar la menor cantidad de becas posibles, pues éstas reducen el ingreso. Si el programa tiene la opción de decidir entre darle beca a un individuo con puntaje inferior a 750 o no, decidirá no hacerlo pues esto le genera mayor ingreso. Por esta razón, no hay problema en que la restricción permita darle beca con un puntaje menor a 750 ya que, gracias a la función objetivo, esto no ocurrirá.

i. Naturaleza de las variables:

$$X_i, Y_i \in \{0,1\}$$

Función objetivo: (1 pto)

$$\max \left\{ \sum_{i=1}^n X_i * A - \sum_{i=1}^n Y_i * R_i * 0,75 * A - \sum_{i=1}^n Y_i * (1 - R_i) * 0,5 * A \right\}$$

2. Se puede eliminar la restricción i: ingreso $> 1,2 * \text{costos}$, pues al maximizar el ingreso podemos ver si se da o no esta restricción. También se puede eliminar la restricción a: relación entre variables, pues dar becas disminuye la función objetivo, lo que hace que no tenga sentido asignar becas a postulantes que no hayan sido aceptados.
3. Nota: Lo que sigue tiene puntaje binario, es decir, todo o nada. El alumno puede tener máximo 1 punto, aún si contesta ambas alternativas planteadas.

Alternativa 1 (1 pto): Al maximizar el ingreso se favorece a los postulantes de menor puntaje, ya que se intentará dar la menor cantidad de becas posibles.

Alternativa 2 (0,5 pts): La restricción h obliga a que un alumno con más de 750 y que ha sido aceptado tenga beca. Por lo tanto, si se nos acaban las B becas, los alumnos restantes con más de 750 puntos no serán aceptados en la universidad. No se les da la posibilidad de pagar el arancel completo, simplemente se rechazan.

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