Pontificia Universidad Católica de Chile
Escuela de Ingeniería
Departamento de Ingeniería Industrial y de Sistemas
Primer Semestre de 2011

ICS 1113, Optimización

Interrogación 3. SOLUCION

Parte 1[25 pts.] : Considere el siguiente problema de optimización no lineal:

$$P) \quad \text{mín} \quad 2(x_1-1)^4 + (x_2-1)^4$$

a) [3 pts.] Muestre que la función objetivo de ese problema es convexa (Use cualquiera de los argumentos basados en los temas vistos en el curso hasta ahora).

Respuesta: Un argumento puede ser el siguiente: se sabe que x^4 es una función convexa. La función objetivo es combinación de x^4 con una función lineal $(x-1)$ y después multiuplicada por un escalar > 0 y eso es convexo. Además es la suma de dos funciones convexas, una en x_1 y la otra en x_2.

Otro argumento puede usar la ciondición de segundo orden diferenciable ya que si f es la función objetivo del problema,

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 24(x_1-1)^2 & 0 \\ 0 & 12(x_2-1)^2 \end{bmatrix}$$

Esta matriz es definida positiva para todo punto $(x_1, x_2) \neq (1, 1)$. El problema es que en $(1, 1)$ es sólo semidefinida positiva, pero es aceptable decir que la convexidad estricta tienen que mantenertse en $(1, 1)$ por continuidad.

b) [5 pts.] La solución óptima del problema es evidente a la vista, pero aún así calcúlela y verifique que cumple la condición de optimalidad de primer orden. ¿Puede usar la condición de segundo orden para confirmar en forma suficiente que efectivamente es un mínimo? Explique.

Respuesta: El gradiente de f es:

$$\nabla f(x_1, x_2) = \begin{bmatrix} 8(x_1-1)^3 \\ 4(x_2-1)^3 \end{bmatrix}$$

Igualando a $(0, 0)$ tenemos que se obtiene $x_1^*, x_2^*) = (1, 1)$. Para la condición de segundo orden tenemos que el Hessiano es:

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 24(x_1-1)^2 & 0 \\ 0 & 12(x_2-1)^2 \end{bmatrix}$$

y en $(1, 1)$ es idénticamente nulo. Luego, no podemos usar la condición de segundo orden ya que se debería tener que ese Heassiano debe ser definido positivo y no lo es.

c) [7 pts.] A partir del punto $(2,2)$ realice 1 iteración del método del gradiente aplicado a este problema. Sea detallado en todos sus pasos y sea ordenado para no equivocarse en los cálculos.

Respuesta: El punto de partida es $(2,2)$. Tenemos que

$$\nabla f(x_1, x_2) = \begin{bmatrix} 8(x_1-1)^3 \\ 4(x_2-1)^3 \end{bmatrix}$$

Luego,

$$\nabla f(2,2) = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

y, entonces la dirección del métodos del gradiente es $h = (h_1, h_2) = (-8, -4)$ Ahora tennemos que resolver el problema en una dimensión:

$$\min_{t \geq 0} f(2 + th_1, 2 + th_2) = 2(2 - 8t - 1)^4 + (2 - 4t - 1)^4 = 2(1 - 8t)^4 + (1 - 4t)^4$$

Para calcular este mínimo en t derivamos e igualamos a 0, lo que nos da:

$$-8(1 - 8t)^3 - 4(1 - 4t)^3 =$$

de donde se deduce que

$$-\sqrt[3]{2}(1 - 8t) = 1 - 4t$$

de donde se obtiene que $t = 0,16$. El nuevo punto calculado por el método es, entonces:

$$\begin{pmatrix} 2 \\ 2 \end{pmatrix} - (0,16) \times \begin{pmatrix} 8 \\ 4 \end{pmatrix} = \begin{pmatrix} 0,72 \\ 1,36 \end{pmatrix}$$

Y hasta aquí llegan las iteraciones ya que era sólo una.

d) [7 pts.] A partir del punto $(2,2)$ realice 1 iteración del método de Newton aplicado a este problema (hágalo minimizando la función restringida a la dirección de Newton, tal como lo vimos en clases). Sea detallado en todos sus pasos.

Respuesta: Vamos a usar parate de la información de la parte anterior:

$$\nabla f(2,2) = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

Además:

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 24(x_1-1)^2 & 0 \\ 0 & 12(x_2-1)^2 \end{bmatrix}$$

luego,

$$\nabla^2 f(2,2) = \begin{bmatrix} 24 & 0 \\ 0 & 12 \end{bmatrix} \Rightarrow [\nabla^2 f(2,2)]^{-1} = \begin{bmatrix} 1/24 & 0 \\ 0 & 1/12 \end{bmatrix}$$

Luego, la dirección de Newton es:

$$h = -[\nabla^2 f(2,2)]^{-1} \nabla f(2,2) = - \begin{bmatrix} 1/24 & 0 \\ 0 & 1/12 \end{bmatrix} \begin{bmatrix} 8 \\ 4 \end{bmatrix} = - \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix}$$

Ahora tenemos que escribir el problema de minimización unidireccional:

$$\min_{t \geq 0} f(2 + th_1, 2 + th_2) = 2(2 - (1/3)t - 1)^4 + (2 - (1/3)t - 1)^4 = 2(1 - (1/3)t)^4 + (1 - (1/3)t)^4$$

y es inmediato ver que esto se minimiza en $t = 3$. Luego, el nuevo punto es:

$$\begin{pmatrix} 2 \\ 2 \end{pmatrix} - (3) \times \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

que es precisamente la solución óptima del problema.

e) [3 pts.] ¿Cómo se compara la solución que usted obtiene en c) con la que obtiene en d)? ¿Cómo se comparan las dos direcciones en que se mueven cada uno de los métodos? Explique con precisión por qué pasa lo que está observando en sus iteraciones.

Respuesta: Podemos ver que el Método de Newton llega en una iteración a la solución óptima del problema, mientras que el método del gradiente no. En clases vimos que para problemas cuadráticos Newton es "instantáneo", y aunque este problema no es cuadrático, ocurre lo mismo. La razón es que la multiplicación por la inversa del Hessiano cambia la orientación de la dirección de descenso y la deja "apuntando" justo hacia el óptimo. Esa es la gran propieda del Método de Newton.

Parte 2[25 pts.] : Las siguientes preguntas de esta parte todas requieren de respuestas breves. Sea preciso.

a) [5 pts.] Las condiciones de KKT para un problema convexo son siempre necesarias y suficientes. ¿Es verdad o no? Explique.

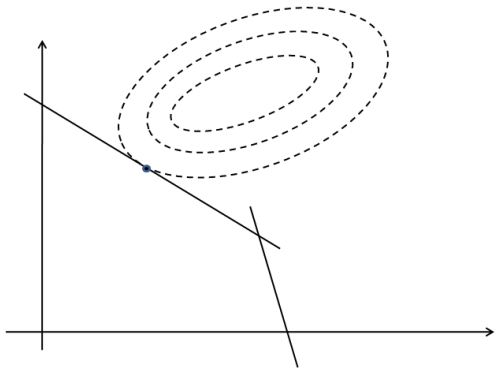
Respuesta: Es falso, las condiciones son, en general, sólo necesarias. Para problemas convexos que cumplan la condición de regularidad de Slater-Uzawa son necesarias y suficientes.

b) [5 pts.] Considere un problema de optimización de la siguiente forma:

$$P) \quad \begin{array}{rl} z^* = \text{mín} & f(x) \\ s.a. & Ax = b, x \geq 0 \end{array}$$

donde f es una función no lineal convexa. ¿Es verdad que una solución óptima del problema P) (en caso de existir) siempre está en un vértice del poliedro definido por las restricciones lineales? Explique con precisión su argumento.

Respuesta: No tiene por qué ser así, la no linealidad de la función objetivo puede hacer que el óptimo esté en el interior de una de las caras de la región factible, como muestra la figura:



c) [5 pts.] Considere el problema de Programación Lineal

$$P) \quad \begin{array}{rl} z^* = \text{mín} & c^T x \\ s.a. & Ax = b, x \geq 0 \end{array}$$

donde A es de m filas y n columnas, $m < n$ y tiene rango completo. Escriba las condiciones de KKT para este problema y muestre que los multiplicadores de Lagrange en el óptimo son las variables duales óptimas del problema P).

Respuesta: Esto fue hecho en ayudantía. Sea λ un vector de m multiplciadores de Lagrange, el Lagrangeano es:

$$L(x, \lambda) = c^T x + \lambda^T (b - Ax)$$

Luego, las condiciones KKT dicen que:

$$\nabla_x L(x,\lambda) = c - A^T\lambda \geq 0, x^T(c - A^T\lambda) = 0.$$

Tenemos que la primera condición es lo mismo que

$$A^T\lambda \leq c$$

luego, λ es solución factible del dual de P). Pero además, de la segunda condición tenemos que

$$c^T x = x^T A^T \lambda = b^T \lambda$$

(donde la segunda ecuación sale del hecho que $Ax = b$). Luego, λ alcanza el mismo valor de función objetivo que el valor óptimo del primal P). Se concluye que λ debe ser óptimo.

d) [5 pts.] Considere un problema de optimización no lineal de la forma:

$$P) \quad \begin{array}{rl} z^* = \text{máx} & f(x) \\ s.a. & g_i(x) \leq b_i, i = 1, \ldots, m \\ & x_j \geq 0, j = 1, \ldots, n \end{array}$$

Escriba las condiciones de KKT para este problema.

Respuesta: Lss condiciones dicen que existen $\lambda_1 \geq 0, \ldots, \lambda_m \geq 0$ tales que:

$$\begin{aligned} \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) &\geq 0 \\ x_j \left(\frac{\partial f}{\partial x_j}(x) + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j}(x) \right) &= 0, j = 1, \ldots, n \\ \lambda_i (g_i(x) - b_i &= 0, i = 1, \ldots, n \\ x_j &\geq 0, j = 1, \ldots, n \end{aligned}$$

e) [5 pts.] Considere las siguientes dos funciones en dos variables:

$$f_1(x,y) = 3x^2 - 2xy + y^2, \quad f_2(x,y) = 2x^2 + e^{2y^2} - 7y^2$$

Indique cuál de esas dos funciones no es convexa y justifíquelo adecuadamente utilizando la caracterización de segundo orden de la convexidad.

Respuesta: Usaremos la condición de segunda orden para la convexidad. Para f_1 tenemos:

$$\nabla^2 f_1(x,y) = \begin{pmatrix} 6 & -2 \\ -2 & 2 \end{pmatrix}$$

y es fácil verificar que esta matriz es definida positiva. Luego, f_1 es convexa. Para f_2 tenemos:

$$\nabla^2 f_2(x,y) = \begin{pmatrix} 4 & 0 \\ 0 & -14 + e^{2y^2} + 16y^2 e^{2y^2} \end{pmatrix}$$

Si evaluamos en $(x, y) = (0, 0)$ tenemos que:

$$\nabla^2 f_2(x, y) = \begin{pmatrix} 4 & 0 \\ 0 & -14 + 4 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & -10 \end{pmatrix}$$

y esta matriz no es semidefinida positiva. Se concluye que f_2 no puede ser convexa.

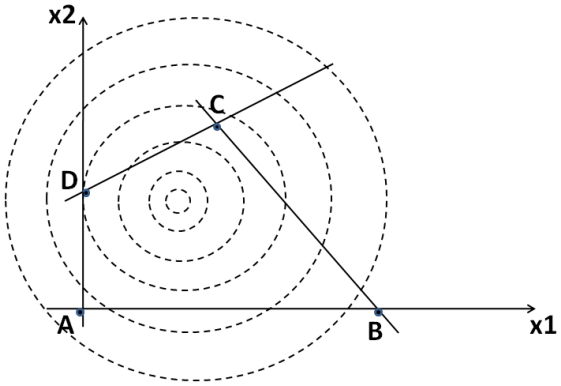
Parte 3[25 pts.] :

1. [13 pts.] Considere el siguiente problema de Optimización no lineal:

$$P) \quad \begin{array}{ll} \text{máx} & (x_1-1)^2+(x_2-1)^2 \\ s.a. & x_1+x_2 \leq 4 \\ & -x_1+x_2 \leq 1 \\ & x_1 \geq 0, x_2 \geq 0 \end{array}$$

a) [4 pts.] Mediante un análisis gráfico determine todos los máximos locales del problema (ind: fíjese en que el problema es de maximización, y haga bien el dibujo).

Respuesta: La figura que sigue muestra la región factible y algunas curvas de nivel de la función objetivo. Mientras más grande sea el círculo, corresponde a un mayor valor (el dibujo no está a escala).



Como el problema es de maximización, se ve claramente que los posibles máximos deben estar en la región factible lo más alejados posible del centro de los círculos. Esta da cuatro máximos locales en los cuatro vértices que se muestran en la figura. LAs coordenadas de estos puntos son: $A = (0,0)$, $B = (4,0)$, $C = (3/2, 5/2)$ y $D = (0,1)$. El fácil ver que el máximo global es B, pero hay que analizar los cuatro puntos.

b) [9 pts.] Para cada uno de los puntos determinados por usted en a), verifique que esos puntos cumplen efectivamente las condiciones KKT y que son puntos regulares.

Respuesta: Comencemos diciendo que la regularidad de los cuatro puntos se sigue simplemente del hecho que las restricciones son lineales y el problema es factible (COndición de Slater-Uzawa aplicada al caso lineal).

En lo que sigue, llamaremos f a la función objetivo, g_1 a la primera restricción, y g_2 a la segunda restricción. El lagrangeano del problema es:

$$L(x_1, x_2, \lambda_1, \lambda_2) = (x_1-1)^2+(x_2-1)^2+\lambda_1(x_1+x_2-4)+\lambda_2(-x_1+x_2-1)$$

Los siguientes son los gradientes de cada una de las funciones involucradas.

$$\nabla f(x_1,x_2) = \begin{pmatrix} 2(x_1-1) \\ 2(x_2-1) \end{pmatrix}, \nabla g_1(x_1,x_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \nabla g_2(x_1,x_2) = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

Ahora mostramos el análisis de las condiciones KKT para cada uno de los cuatro puntos. Recordemos que el problema es de maximización, pero podemos manejar implícitamente esa característica tomando $-f$ y pensando en el problema como de minimización.

Para el punto $A = (0, 0)$, dado que las dos restricciónes son inactivas podemos asegurar de inmediato que $\lambda_1 = \lambda_2 = 0$ (Por holguras complementarias). Por otro lado,

$$\nabla f(0,0) = \begin{pmatrix} -2 \\ -2 \end{pmatrix},$$

luego $-\nabla f(0,0) \geq 0$, como debe efectivamente verificarse, además de que obviamente $x^T \nabla f(x) = 0$. Luego, A cumple KKT.

Para el punto B tenemos que la segunda restricción es no activa. Luego, $\lambda_2 = 0$. Tenemos:

$$\nabla f(4,0) = \begin{pmatrix} 6 \\ -2 \end{pmatrix},$$

Debemos verificar que existe $\lambda_1 \geq 0$ tal que

$$-\nabla f(x) + \lambda_1 \nabla g_1(x) \geq 0,\ x^T(-\nabla f(x) + \lambda_1 \nabla g_1(x)) = 0$$

Una de las coordenadas es > 0, luego, se debe tener igualdad a cero en la correspondiente coordenada del gradiente del lagrangeano:

$$\begin{pmatrix} -6 \\ 2 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \geq 0$$

con $-6 + \lambda_1 = 0$ ya que $x_1 = 4 > 0$. Esto implica que $\lambda_1 = 6$ y este valor satisface además que $2 + \lambda_1 = 2 + 6 = 8 > 0$. Luego, se cumple KKT para este punto.

Para el punto C tenemos que las dos restricciones son activas, pero además el punto tiene coordenadas > 0, así que basta aplicar la versión "simple" de KKT. Tenemos:

$$\nabla f(3/2, 5/2) = \begin{pmatrix} -1 \\ 3 \end{pmatrix},$$

Tenemos que encontrar $\lambda_1 \geq 0$ y $\lambda_2 \geq 0$ tales que:

$$-\nabla f(x) + \lambda_1 \nabla g_1(x) + \nabla_2 \nabla g_2(x) = 0$$

Es decir,

$$\begin{pmatrix} 1 \\ -3 \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 0$$

Y esto se resuelve con $\lambda_1 = 1, \lambda_2 = 2$. Luego, se cumplen las condiciones KKT.

Finalmente, para el punto D tenemos que la restricción 1 es inactiva. Luego, $\lambda_1 = 0$. Tenemos que se debe encontrar $\lambda_2 \geq 0$ tal que

$$-\nabla f(x) + \lambda_2 \nabla g_2(x) \geq 0,\ x^T(-\nabla f(x) + \lambda_2 \nabla g_2(x)) = 0$$

con igualdad a 0 en la segunda coordenada ya que $x_2 = 1 > 0$. Es decir:

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} \geq 0$$

con $0 + \lambda_2 = 0$ ya que $x_2 = 1 > 0$. Esto implica que $\lambda_2 = 0$ y este valor satisface además que $2 + \lambda_2 \times -1 = 2 > 0$. Luego, se cumple KKT para este punto.

2. [12 pts.] Considere el siguiente problema de Optimización no lineal:

$$P) \quad \begin{array}{ll} \text{máx} & (x_1)^{\alpha_1}(x_2)^{\alpha_2} \dots (x_n)^{\alpha_n} \\ s.a. & \sum_{i=1}^{n} x_i = 1 \\ & x_i \geq 0, i = 1, \dots, n \end{array}$$

donde $\alpha_i > 0, i = 1, \dots, n$.

a) [3 pts.] Argumente que no es necesario consider las restricciones $x_i \geq 0, i = 1, \dots, n$

Respuesta: Es obvio de observar que la función objetivo maximiza y si algún $x_j = 0$, entonces vale 0, y debiera ser mayor que eso.

b) [6 pts.] Use las condiciones de optimalidad para problemas restringidos para concluir que si $x_1^*, \dots, x_n^*$ es solución óptima del problema P), entonces

$$\frac{x_i^*}{x_j^*} = \frac{\alpha_i}{\alpha_j}, i = 1, \dots, n; j = 1, \dots, n.$$

Respuesta: Es conveniente considerar el problema equivalente:

$$P) \quad \begin{array}{ll} \text{máx} & \sum_{i=1}^{n} \alpha_i \log(x_i) \\ s.a. & \sum_{i=1}^{n} x_i = 1 \end{array}$$

El Lagrangeano del problema es (sin considerar $x \geq 0$):

$$L(x, \lambda) = \sum_{i=1}^{n} \alpha_i \log(x_i) + \lambda(\sum_{i=1}^{n} x_i - 1)$$

Las condiciones KKT (o Lagrange) nos dicen que:

$$\alpha_j \frac{1}{x_j} + \lambda = 0, j = 1, \dots n$$

de donde se deduce, para $i \neq j$ que

$$\alpha_i \frac{1}{x_i} = \alpha_j \frac{1}{x_j}$$

de donde se tiene el resultado.

c) [3 pts.] Usando lo anterior, determine la solución óptima del problema.

Respuesta: Usando el resultado anterior tenemos que:

$$\frac{x_i}{x_1} = \frac{\alpha_i}{\alpha_1}, i = 2, \ldots, n$$

Es decir,

$$x_i = \frac{\alpha_i}{\alpha_1} x_1$$

Reemplazando en la restricción tenemos que:

$$\sum_{i=1}^{n} x_i = 1 \Leftrightarrow x_1 + \sum_{i=2}^{n} x_i = x_1 + \sum_{i=2}^{n} \frac{\alpha_i}{\alpha_1} x_1 = 1$$

Luego,

$$x_1 = \left(1 + \frac{1}{\alpha_i} \sum_{i=2}^{n} \alpha_i\right)^{-1}$$

y esto determina todos los valores.

Parte 4[25 pts.] :
Un pais quiere mejorar la calidad de su sistema educacional y para eso ha desarrollado un programa de formación de profesores para los próximos T años. En este programa, año a año, se forman una cierta cantidad de profesores "normales", y además, un porcentaje de estos recibe educación y entrenamiento adicional para transformase en "profesores de excelencia". Los profesores de excelencia, a su vez, participan en la formación de nuevos profesores "normales". Cada año, por cada profesor de excelencia, se forman γ profesores nuevos, $\gamma > 0$. Adicionalmente, cada año, un procentaje δ del total de todos los profesores (normales y de excelencia) abandona su carrera (por fallecimiento, retiro o cambio de carrera). El estado, mediante incentivos e inversión en programas adicionales de formación, pueda hacer que en el año t un porcentaje u_t, $0 \leq u_t \leq 100, t = 1, \ldots, T$, de los nuevos profesores se transformen en profesores de excelencia. Desde luego, a mayor porcentaje u_t, hay mayores costos y existen presupuestos limitados a nivel de la nación. Específicamente, por cada punto porcentual de profesores que se transforman en profesores de excelencia en el año t, se invierten $\$C_t$. El presupuesto para formación de profesores de excelencia sobre los T años que ha aprobado el legistaltivo de este país es igual a $\$K$. Asuma que inicialmente existe una cantidad x_0 de profesores "normales" y que la cantidad inicial de profesores de excelencia es igual a cero.

1. [10 pts.] Formule un modelo de optimización que permita determinar qué porcentaje de los profesores "normales" de cada año debe ser transformado en profesores de excelencia de modo tal que se maximice la cantidad de profesores de excelencia *al final del horizonte del programa*, es decir, en el año T, y que no se exceda el presupuesto total asignado al programa.

   Respuesta: Definamos las siguientes variables:

   x_t**:** la cantidad de profesores en el año t.

   y_t**:** la cantidad de profesores de excelencia en el año t.

   El modelo es:

$$\begin{array}{rl} \text{máx} & y_T \\ s.a. & x_t = (1-\delta)x_{t-1} + \gamma y_t, t = 1, \ldots, T \\ & y_{t+1} = y_t + (u_t/100)x_t, t = 1, \ldots, T \\ & \sum_{t=1}^{T} C_t u_t \leq K \\ & x_t \geq 0, y_t \geq 0, 0 \leq u_t \leq 100, t = 1, \ldots, T \end{array}$$

2. [15 pts.] Escriba el código necesario en AMPL que corresponde su modelo anterior (lo que escribiría en el archivo .MOD). Defínase los conjuntos, parámetros y variables que necesite pero sea claro en indicar qué es cada cosa, ojalá en forma clara con respecto a su modelo de a).