PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
Departamento de Ingeniería Industrial y de Sistemas

Curso: ICS1113-Optimización
Semestre: 02-2016
Profesores: S. Encina - R. Cuadrado - F. Palomino
A. Cataldo - A. Lorca - J.T. Marquinez
G. Angulo

Solución Tarea 3

Fecha de Entrega: 25 de noviembre.

Reglas de la Tarea

No seguir estas reglas generará una penalización en la nota de la tarea.

- La Tarea se desarrolla en forma grupal y estos grupos **deben** ser los mismos grupos ya asignados para el proyecto semestral. Cada grupo debe entregar una sola tarea.
- Se debe contestar en hojas **independientes** cada una de las cuatro preguntas de la Tarea. Estas hojas deben ser **blancas y de tamaño carta**, y está a su decisión si escribirla en computador o a mano, mientras **esté ordenado**. En cada hoja debe colocar su **número de grupo**[1] **con letra clara y legible**. Las hojas de cada pregunta deben estar corcheteadas, entregando así cuatro grupos de hojas el día de la entrega.
- El plazo de entrega vence impostergablemente el día viernes 25 de noviembre a las 12:30 horas puntualmente. Aquellas Tareas no entregadas en la fecha y hora indicadas, serán consideradas como **Tarea NO Entregada**.
- La Tarea se entrega en la **secretaría del segundo piso del edificio Raúl Devés**. Deben entregar las preguntas por separado en el buzón que corresponde a cada pregunta.
- Esta tarea es grupal y el desarrollo y discusión debe ocurrir dentro de cada grupo. No se distribuyan la resolución de las preguntas por separado, hagan realmente un trabajo grupal de desarrollo ya que el no hacerlo va contra la idea de aprendizaje colaborativo. Pueden discutir los problemas con los profesores y los ayudantes del curso, pero al final cada grupo debe entregar sus propias soluciones, desarrolladas y escritas por el grupo. La copia o intento de copia a otros grupos será sancionada dependiendo la gravedad (a ser determinada por el equipo docente del curso) con consecuencias que podrían ir desde un 1.0 en la nota de la Tarea hasta la escalación a la Dirección de Docencia de la Escuela, con posible **REPROBACIÓN AUTOMÁTICA** del curso[2].
- Cualquier duda sobre el enunciado, enviarlas al correo electrónico `dudas.tareas.ics1113@gmail.com`. Este correo centralizará todas las dudas (únicamente) sobre el **enunciado de la tarea**. Cualquier correo enviado a otra dirección respecto a la tarea, no será considerado (y por lo tanto, no será respondido). Tampoco lo serán aquellos correos que lleguen a este buzón que contengan preguntas no relacionadas con la tarea. Este correo será revisado desde la publicación de este enunciado hasta el día antes de la entrega, dos veces al día (una en la mañana y otra en la tarde-noche).

[1] Deben escribir «Grupo XX», donde XX corresponde al número de grupo.
[2] Tener en consideración, para esto, el Código de Honor de la Escuela

Problema 1. (15 puntos)

a) (3 puntos) De una intuición usando sus palabras de por qué el algoritmo de Dijkstra termina y no se queda en *loops* infinitos.

b) (7 puntos) Considera la red de la figura 1,

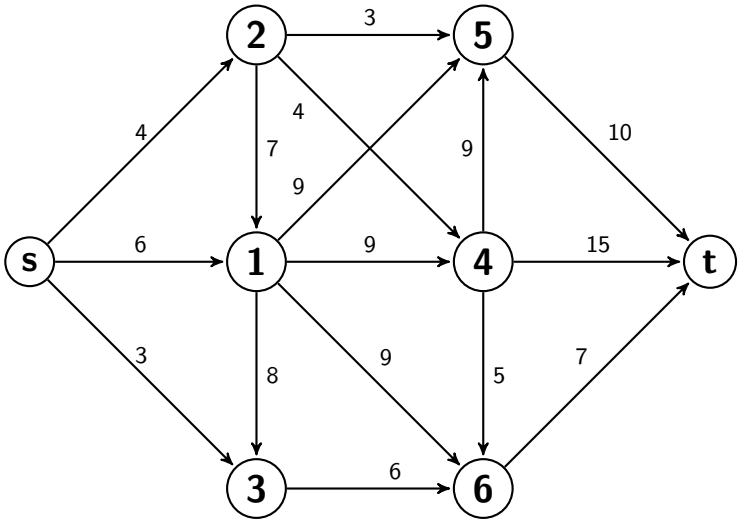


Figura 1: Red con costos.

donde los números que acompañan los arcos son las distancias de pasar por ahí.

i) (3 puntos) Resuelva el problema de encontrar el camino más corto entre el nodo s y el nodo t usando Dijkstra.

ii) (1 puntos) ¿Cuál es el camino más corto entre 2 y t?

iii) (3 puntos) Construya el árbol que contenga el camino más corto entre cualquier nodo y t. Explique claramente cómo lo calcula.

c) (5 puntos) Considere la red que se muestra en la figura 2,

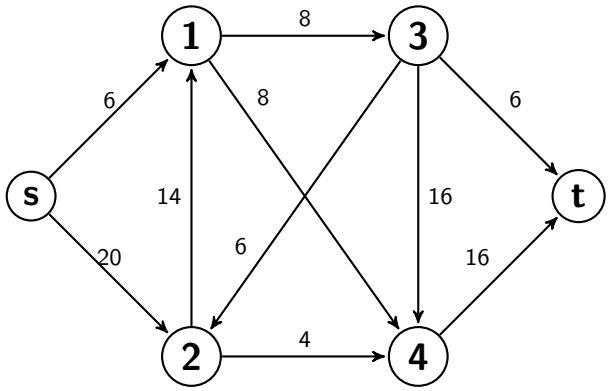


Figura 2: Red con capacidades máximas.

donde los números que acompañan a los arcos son las capacidades máximas de flujo. Se le pide que encuentre el corte mínimo asociada a la red.

Solución Problema 1.

a) El algoritmo Dijkstra va ingresando un nodo a la vez y no vuelve a ingresarlo de nuevo, en el peor de los casos revisamos una vez cada nodo, por lo que termina y como no vuelve a revisar los nodos que ya ingresaron a V, entonces no puede caer en *loops*.

b) *i*) Al utilizar el algoritmo de Dijkstra, llegamos a la siguiente tabla resumen:

it.	s	1	2	3	4	5	6	t	V
0	$(0,s)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$\{\}$
1	$(0,s)$	$(6,s)$	$(4,s)$	$(3,s)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$\{s\}$
2	–	$(6,s)$	$(4,s)$	$(3,s)$	$(-,\infty)$	$(-,\infty)$	$(9,3)$	$(-,\infty)$	$\{s,3\}$
3	–	$(6,s)$	$(4,s)$	–	$(8,2)$	$(7,2)$	$(9,3)$	$(-,\infty)$	$\{s,3,2\}$
4	–	$(6,s)$	–	–	$(8,2)$	$(7,2)$	$(9,3)$	$(-,\infty)$	$\{s,3,2,6\}$
5	–	–	–	–	$(8,2)$	$(7,2)$	$(9,3)$	$(17,5)$	$\{s,3,2,1,5\}$
6	–	–	–	–	$(8,2)$	–	$(9,3)$	$(17,5)$	$\{s,3,2,1,5,4\}$
7	–	–	–	–	–	–	$(9,3)$	$(16,6)$	$\{s,3,2,1,5,4,6\}$
8	–	–	–	–	–	–	–	$(16,6)$	$\{s,3,2,1,5,4,6,t\}$

Entonces, el camino mas corto es $s-3-6-t$ con un costo de 16.

ii) Como 2 no está en el camino, no podemos usar directamente *i*) para responder esto. Pero podemos responderla con la respuesta de *iii*).

iii) Para esta pregunta, hacemos un grafo que de vuelta los arcos como se ve en la figura 4 y usamos el algoritmo de Dijkstra partiendo de t. Luego, el arbol que nos queda le volvemos a dar vuelta los nodos y estamos listos.

it.	s	1	2	3	4	5	6	t	V
0	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(0,t)$	$\{\}$
1	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(-,\infty)$	$(15,t)$	$(10,t)$	$(7,t)$	$(0,t)$	$\{t\}$
1	$(-,\infty)$	$(16,6)$	$(-,\infty)$	$(13,6)$	$(12,6)$	$(10,t)$	$(7,t)$	–	$\{t,6\}$
1	$(-,\infty)$	$(16,6)$	$(13,5)$	$(13,6)$	$(12,6)$	$(10,t)$	–	–	$\{t,6,5\}$
1	$(-,\infty)$	$(16,6)$	$(13,5)$	$(13,6)$	$(12,6)$	–	–	–	$\{t,6,5,4\}$
1	$(16,3)$	$(16,6)$	$(13,5)$	$(13,6)$	–	–	–	–	$\{t,6,5,4,3\}$
1	$(16,3)$	$(16,6)$	$(13,5)$	–	–	–	–	–	$\{t,6,5,4,3,2\}$
1	$(16,3)$	$(16,6)$	–	–	–	–	–	–	$\{t,6,5,4,3,2,1\}$
1	$(16,3)$	–	–	–	–	–	–	–	$\{t,6,5,4,3,2,1,s\}$

Luego, el árbol pedido esta en la figura 3

c) Para encontrar el corte mínimo podemos utilizar el algoritmo de Ford-Fulkerson para encontrar el flujo máximo y luego con eso ver el corte mínimo. Este procedimiento es escalable cuando hay mucho nodos, ya que se pueden enumerar los cortes también para resolver el problema, pero eso en grandes instancias no es muy eficiente.

Primero debemos encontrar un st-camino para ir empujando flujo. En primera instancia puede ser cualquiera, por eso elegimos el camino $s-2-1-3-t$, en el cual se puede aumentar el flujo f en 6. Con esto, el grafo residual que nos queda es el que se ve en la figura 6.

En la segunda iteración podemos tomar el camino $s-2-1-4-t$ que puede llevar a lo mas $\Delta f = 8$ desde s a t. Con lo que nos queda el grafo residual que se ve en la figura 7.

Luego, viendo el grafo residual acual, podemos tomar el camino $s-2-4-t$, donde $\Delta f = 4$, al pasar ese flujo, el grafo residual es el que se ve en la figura 8. Finalmente, al empujar flujo por el camino $s-1-3-4-t$ ($\Delta f = 2$), podemos ver que en el grafo residual dado en la figura 9 no nos quedan caminos que partan en s y terminen en t, en particular, no podemos pasar por los arcos del medio, ya que todos apuntan hacia el conjunto conexo que contiene a s. Entonces, podemos ver que el corte mínimo es el dado por el conjunto $S = \{s, 1, 2\}$. Con un valor de 20 al igual que el flujo máximo.

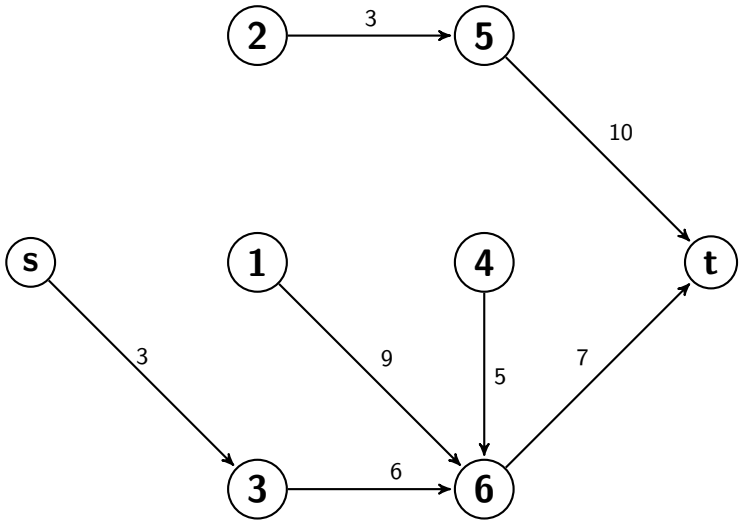


Figura 3: Árbol de caminos mas cortos hacia t.

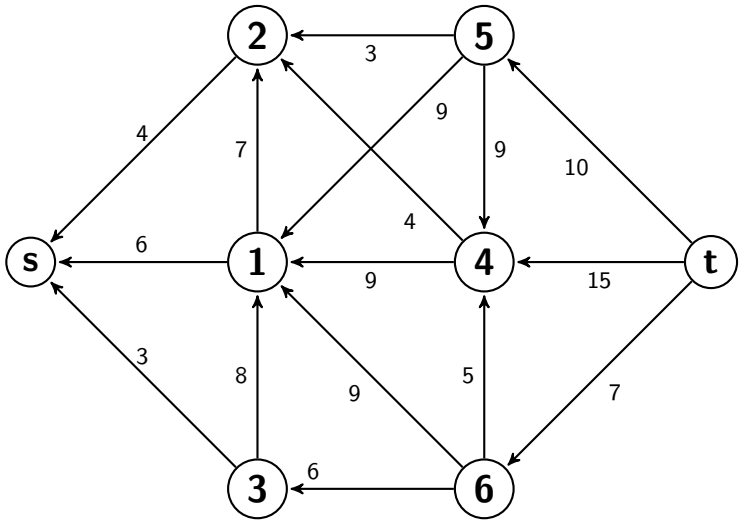


Figura 4: Grafo con arcos al revés.

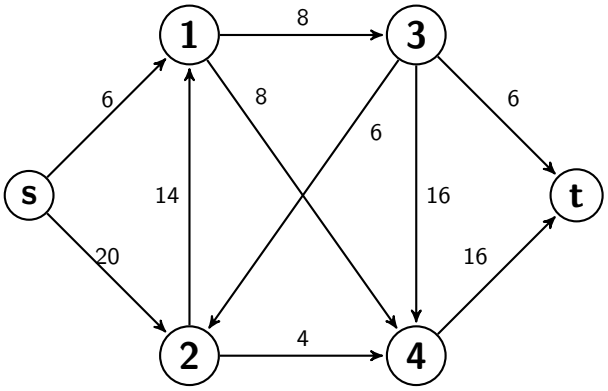


Figura 5: grafo residual inicial

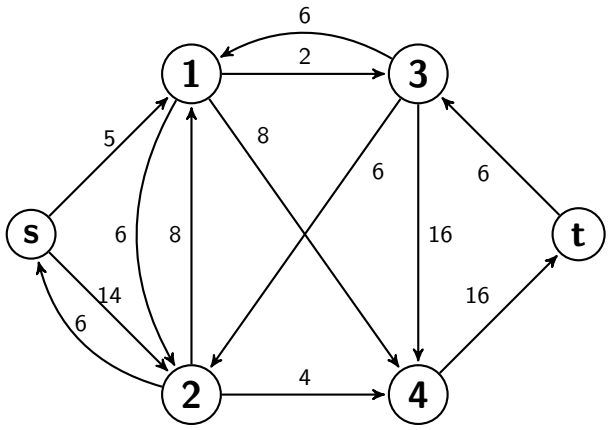


Figura 6: grafo residual luego de la primera iteración

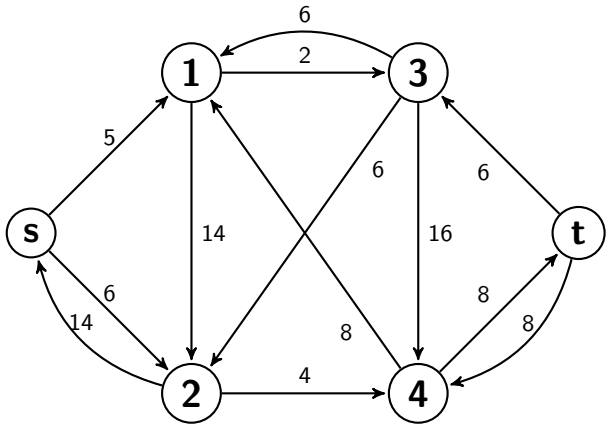


Figura 7: grafo residual luego de la segunda iteración

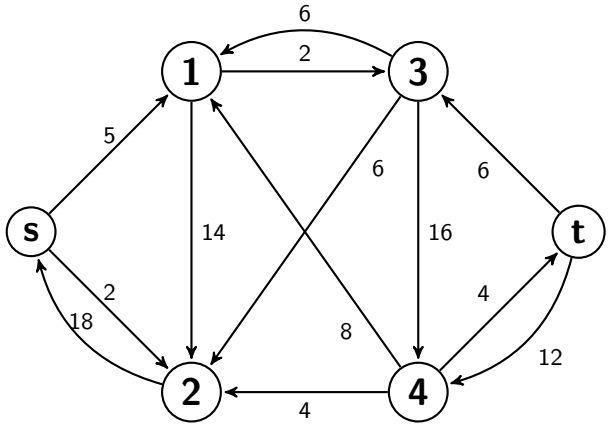


Figura 8: grafo residual luego de la tercera iteración

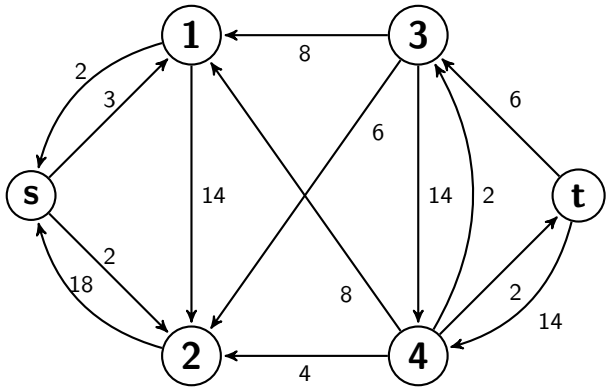


Figura 9: grafo residual luego de la cuarta iteración

Problema 2. (15 puntos)

a) (7 puntos) Queremos resolver un problema de la forma $\min_{x\in\mathbb{R}^3} f(x)$ con $f(x) = \frac{1}{2}x^T Hx + c^T x$. Para esto considere

$$H = \begin{pmatrix} 10 & -18 & 2 \\ -18 & 40 & -1 \\ 2 & -1 & 3 \end{pmatrix} \quad \text{y} \quad c = \begin{pmatrix} 12 \\ -47 \\ -8 \end{pmatrix}.$$

i) (1 punto) Encuentre los puntos críticos del problema, el óptimo y su valor óptimo. Justifíquelo.

ii) (1 punto) Defina un criterio de parada para el método del gradiente que utilice la respuesta de *i*).

iii) (3 puntos) Utilice el método del descenso mas pronunciado para resolver el problema. En la parte de resolver el problema de óptimización que entrega el largo óptimo del paso, puede utilizar excel o ampl para resolverlo. Parta del punto $(0, 0, 0)$ y resuelva con un error de $0,35$ para el criterio que definió en *ii*).

iv) (1 punto) Si se hubiese tomado un error de $0,4$ y $0,5$ cuantas iteraciones se hubiesen realizado en cada una de ellas.

v) (1 punto) Calcule el error de la estimación del óptimo en cada iteración (véase en la Observación) y luego grafique el error en función de las iteraciones. ¿A qué función se parece el gráfico?¿Qué interpretación tiene esto? Comente.

vi) BONUS (2 punto) ¿Se cumple que $\nabla f(x^{(k)})^T \nabla f(x^{(k+1)}) = 0$ para sus iteraciones? En caso de no ser cierto, ¿por qué cree usted que ocurre eso? En caso de ser cierto, deje el calculo numérico para comprobarlo.

OBS: El error en cada iteración lo puedes calcular según la siguiente ecuación:

$$\text{ERROR}_k = ||x^{(k)} - x^*||$$

siendo x^* el óptimo del problema.

b) (8 puntos) En esta pregunta queremos verificar que el método de Newton con paso fijo de 1, no necesariamente converge. De hecho, el método puede diverger si el punto inicial no es cercano al óptimo, cosa que no se puede verificar si es que no sabemos cuál es el óptimo o dónde está más o menos. Para esto consideraremos dos ejemplos:

- ◇ La función $f(x) = \log(e^x + e^{-x})$.
- ◇ La función $g(x) = -\log(x) + x$.

En función de estos ejemplos se le pide:

i) (3 punto) Graficar $f(x)$ y $g(x)$ y encontrar sus puntos críticos y demostrar cuál es el óptimo global.

ii) (3 puntos) Utilice el método de Newton e itere 4 veces (llegar a k=4) para mín $f(x)$ y mín $g(x)$ con paso fijo de 1 y partiendo con $x^{(0)} = 1{,}2$ y $x^{(0)} = 1{,}3$ para $f(x)$ y con los puntos $x^{(0)} = 1{,}1$ para $g(x)$.

iii) (2 puntos) Calcule y grafique, para los tres casos de *ii*), el error (calculando el error como en *a*)) de las iteraciones. ¿Qué puede decir de la convergencia? Comente.

Solución Problema 2.

a) *i*) Para esto, hay que derivar $f(x)$, y como H es una matriz simetrica y recordando los cursos de calculo, tenemos que:

$$\nabla f(x) = Hx + c^T$$
$$\nabla^2 f(x) = H.$$

Además, como veamos que H es una matriz definida positiva. Para esto vemos que los determinantes de sus menores principales son positivos:

$$\det_1 = 10 \quad \det_2 = 400 - 324 = 76,$$

$$\det_3 = |H| = 2(18 - 80) + 1(-10 + 36) + 3(76) = -124 + 26 + 228 = 130.$$

Por lo que tenemos que la matriz H es definida positiva, lo que significa que el problema es convexo (es una parábola en 3D). Ahora busquemos los puntos críticos:

$$\nabla f(x) = Hx + c^T = \vec{0}$$
$$Hx = -c^T$$
$$x = -H^{-1}c^T.$$

Lo anterior, se puede hacer, ya que H es definida positiva, lo que nos indica que el único punto que cumple $\nabla f(x) = 0$ es el punto $x = (4, 3, 1)$. Y como la función es convexa, sabemos que ese punto es un óptimo global.

ii) Como los criterios buscan aproximar o decirnos que tan lejos estamos del óptimo (a menos que no tengamos información o tiempo y en ese caso paramos en una cantidad fija de iteraciones), podemos definir el criterio de $||x - x^*|| \leq \epsilon$, donde ϵ es el error tolerable, x^* el óptimo (calculado en *i*)) y$|||\dot{}|||$ es alguna norma, podemos usar la distancia euclidiana (ojo que dependiendo de esto, la cantidad de iteraciones en *iii*) puede crecer mucho, los alumnos deberían elegir inteligentemente).

iii) Al resolver con los datos especificados, tenemos que las iteraciones nos dan:

$x^{(1)} = (0, 0, 0)$,	$d^{(1)} = (-12, 47, 8)$,	$\lambda^{(1)} = 0{,}0221418$
$x^{(2)} = (-0{,}265702, 1{,}04067, 0{,}177134)$,	$d^{(2)} = (9{,}03472, 0{,}767901, 9{,}04067)$,	$\lambda^{(2)} = 0{,}142798$
$x^{(3)} = (1{,}02443, 1{,}15032, 1{,}46812)$,	$d^{(3)} = (-4{,}47484, 20{,}8952, 2{,}69709)$,	$\lambda^{(3)} = 0{,}0222055$
$x^{(4)} = (0{,}925069, 1{,}61431, 1{,}52801)$,	$d^{(4)} = (4{,}75083, 0{,}606938, 3{,}18014)$,	$\lambda^{(4)} = 0{,}147853$
$x^{(5)} = (1{,}62749, 1{,}70405, 1{,}9982)$,	$d^{(5)} = (-1{,}59851, 10{,}1313, 0{,}454453)$,	$\lambda^{(5)} = 0{,}0224132$
$x^{(6)} = (1{,}59166, 1{,}93112, 2{,}00839)$,	$d^{(6)} = (2{,}82672, 0{,}413586, 0{,}722625)$,	$\lambda^{(6)} = 0{,}161412$
$x^{(7)} = (2{,}04793, 1{,}99788, 2{,}12503)$,	$d^{(7)} = (-0{,}767581, 6{,}07271, -0{,}47307)$,	$\lambda^{(7)} = 0{,}0227508$
$x^{(8)} = (2{,}03047, 2{,}13604, 2{,}11427)$,	$d^{(8)} = (1{,}91544, 0{,}221253, -0{,}267697)$,	$\lambda^{(8)} = 0{,}174851$
$x^{(9)} = (2{,}36538, 2{,}17472, 2{,}06746)$,	$d^{(9)} = (-0{,}643752, 4{,}65547, -0{,}758421)$,	$\lambda^{(9)} = 0{,}0228986$
$x^{(10)} = (2{,}35064, 2{,}28133, 2{,}05009)$,	$d^{(10)} = (1{,}45726, 0{,}108611, -0{,}570234)$,	$\lambda^{(10)} = 0{,}17849$
$x^{(11)} = (2{,}61075, 2{,}30071, 1{,}94831)$,	$d^{(11)} = (-0{,}591292, 3{,}91331, -0{,}765718)$,	$\lambda^{(11)} = 0{,}0229221$
$x^{(12)} = (2{,}59719, 2{,}39041, 1{,}93076)$,	$d^{(12)} = (1{,}19397, 0{,}0637376, -0{,}596254)$,	$\lambda^{(12)} = 0{,}17899$
$x^{(13)} = (2{,}8109, 2{,}40182, 1{,}82404)$,	$d^{(13)} = (-0{,}524318, 3{,}34744, -0{,}692093)$,	$\lambda^{(13)} = 0{,}022925$
$x^{(14)} = (2{,}79888, 2{,}47856, 1{,}80817)$,	$d^{(14)} = (1{,}00893, 0{,}0456163, -0{,}543714)$,	$\lambda^{(14)} = 0{,}179048$
$x^{(15)} = (2{,}97953, 2{,}48673, 1{,}71082)$,	$d^{(15)} = (-0{,}455827, 2{,}87322, -0{,}604788)$,	$\lambda^{(15)} = 0{,}0229253$
$x^{(16)} = (2{,}96908, 2{,}5526, 1{,}69695)$,	$d^{(16)} = (0{,}862053, 0{,}0364786, -0{,}476424)$,	$\lambda^{(16)} = 0{,}179055$
$x^{(17)} = (3{,}12344, 2{,}55913, 1{,}61165)$,	$d^{(17)} = (-0{,}393317, 2{,}4683, -0{,}522684)$,	$\lambda^{(17)} = 0{,}0229253$
$x^{(18)} = (3{,}11442, 2{,}61572, 1{,}59966)$,	$d^{(18)} = (0{,}739377, 0{,}0305483, -0{,}412115)$,	$\lambda^{(18)} = 0{,}179056$
$x^{(19)} = (3{,}24681, 2{,}62119, 1{,}52587)$,	$d^{(19)} = (-0{,}338481, 2{,}12098, -0{,}45005)$,	$\lambda^{(19)} = 0{,}0229253$
$x^{(20)} = (3{,}23905, 2{,}66981, 1{,}51556)$,	$d^{(20)} = (0{,}634987, 0{,}0260183, -0{,}354954)$,	$\lambda^{(20)} = 0{,}179056$
$x^{(21)} = (3{,}35275, 2{,}67447, 1{,}452)$,	$d^{(21)} = (-0{,}291027, 1{,}82268, -0{,}387022)$,	$\lambda^{(21)} = 0{,}0229253$
$x^{(22)} = (3{,}34608, 2{,}71625, 1{,}44313)$,	$d^{(22)} = (0{,}545579, 0{,}0222931, -0{,}305275)$,	$\lambda^{(22)} = 0{,}179059$
$x^{(23)} = (3{,}44377, 2{,}72025, 1{,}38846)$,	$d^{(23)} = (-0{,}250152, 1{,}56639, -0{,}332678)$,	$\lambda^{(23)} = 0{,}0229253$
$x^{(24)} = (3{,}43803, 2{,}75616, 1{,}38084)$,	$d^{(24)} = (0{,}46883, 0{,}0191386, -0{,}262418)$,	$\lambda^{(24)} = 0{,}179059$
$x^{(25)} = (3{,}52198, 2{,}75958, 1{,}33385)$,	$d^{(25)} = (-0{,}214992, 1{,}34614, -0{,}285923)$,	$\lambda^{(25)} = 0{,}0229253$
$x^{(26)} = (3{,}51705, 2{,}79044, 1{,}32729)$,	$d^{(26)} = (0{,}402898, 0{,}0164417, -0{,}22554)$,	$\lambda^{(26)} = 0{,}179059$
$x^{(27)} = (3{,}58919, 2{,}79339, 1{,}28691)$,	$d^{(27)} = (-0{,}184766, 1{,}15686, -0{,}245726)$,	$\lambda^{(27)} = 0{,}0229253$
$x^{(28)} = (3{,}58496, 2{,}81991, 1{,}28128)$,	$d^{(28)} = (0{,}346245, 0{,}0141275, -0{,}193833)$,	$\lambda^{(28)} = 0{,}179058$
$x^{(29)} = (3{,}64696, 2{,}82244, 1{,}24657)$,	$d^{(29)} = (-0{,}158786, 0{,}994196, -0{,}211177)$,	$\lambda^{(29)} = 0{,}0229253$
$x^{(30)} = (3{,}64332, 2{,}84523, 1{,}24173)$,	$d^{(30)} = (0{,}29756, 0{,}0121397, -0{,}16658)$,	$\lambda^{(30)} = 0{,}179055$
$x^{(31)} = (3{,}6966, 2{,}84741, 1{,}2119)$,	$d^{(31)} = (-0{,}136457, 0{,}854401, -0{,}181485)$,	$\lambda^{(31)} = 0{,}0229254$
$x^{(32)} = (3{,}69347, 2{,}86699, 1{,}20774)$,	$d^{(32)} = (0{,}255722, 0{,}010431, -0{,}143159)$,	$\lambda^{(32)} = 0{,}17905$

Cabe destacar, que si colocaban el criterio por ejemplo, la norma a la cuarta, las iteraciones bajaban a 22. Y podían bajar ese criterio para que fuese menos estricto, por eso se les dejo la flexibilidad de dar el criterio, para así no trabajar tanto en esta pregunta. Pero aquí coloque la más larga y estricta.

iv) Para no hacer todo de nuevo, basta con anotar los errores, los cuales se ven en la siguiente lista:

$$\begin{aligned}
\text{error}_1 &= 26{,}000000 \\
\text{error}_2 &= 22{,}712312 \\
\text{error}_3 &= 12{,}494443 \\
\text{error}_4 &= 11{,}654141 \\
\text{error}_5 &= 8{,}304701 \\
\text{error}_6 &= 7{,}959434 \\
\text{error}_7 &= 6{,}080516 \\
\text{error}_8 &= 5{,}867082 \\
\text{error}_9 &= 4{,}492525 \\
\text{error}_{10} &= 4{,}339568 \\
\text{error}_{11} &= 3{,}318318 \\
\text{error}_{12} &= 3{,}205772 \\
\text{error}_{13} &= 2{,}450804 \\
\text{error}_{14} &= 2{,}367718 \\
\text{error}_{15} &= 1{,}810067 \\
\text{error}_{16} &= 1{,}748707 \\
\text{error}_{17} &= 1{,}336843 \\
\text{error}_{18} &= 1{,}291525 \\
\text{error}_{19} &= 0{,}987338 \\
\text{error}_{20} &= 0{,}953868 \\
\text{error}_{21} &= 0{,}729208 \\
\text{error}_{22} &= 0{,}704489 \\
\text{error}_{23} &= 0{,}538562 \\
\text{error}_{24} &= 0{,}520305 \\
\text{error}_{25} &= 0{,}397758 \\
\text{error}_{26} &= 0{,}384274
\end{aligned}$$

Así, se ve claramente que si hubiésemos tomado un error de 0,4, hubiésemos parado en la iteración 24 y con 0,5 en la iteración 24 también.

v) El error fue calculado en la parte anterior, podemos ver el gráfico en la figura 10, pero es difícil ver a que función se parece el error, para esto hacemos un gáfico sacando logaritmo en los ejes, lo que se puede ver en la figura 11. Podemos ver que el error disminuye mas rápido que cualquier polinomio (los polinomios son rectas en este tipo de gráficos). Se puede ver también, que el error disminuye más rápido en un principio, pero a medida que nos aceracmos al óptimo, el error decrece cada vez mas lento.

vi) En el Bonus, depende de como hicieron el problema, pero en general se espera que no les de exactamente 0, esto debido a que hicieron a mano el ejercicio, o que el computador le da los valores aproximados, ya que no le muestra las fracciones o le trunca algunos decimales.

b) *i*) Para encontrar los puntos críticos de las funciones $f(x)$ y $g(x)$ las derivamos e igualamos a 0.

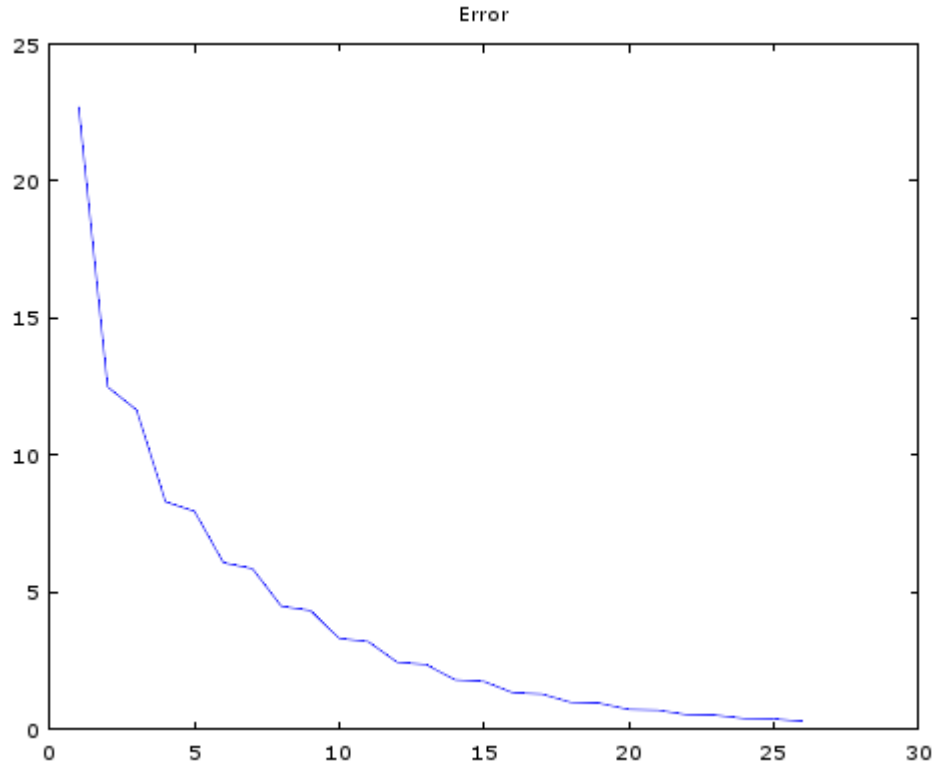


Figura 10: Gráfico del error vs iteraciones

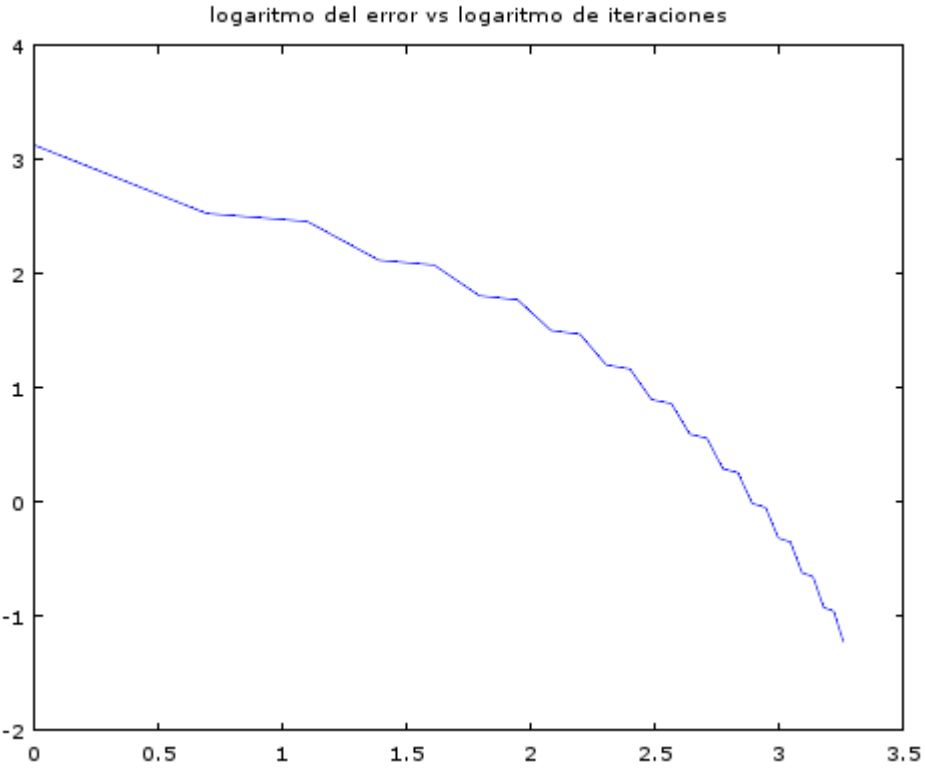


Figura 11: Gráfico del log(error) vs log(iteraciones)

Entonces tenemos que

$$
\begin{aligned}
f(x) &= log(e^x + e^{-x}) \\
f'(x) &= \frac{1}{e^x + e^{-x}}(e^x - e^{-x}) = \frac{e^{2x} - 1}{e^{2x} + 1} \\
f''(x) &= \frac{2e^{2x}(e^{2x} + 1) - 2e^{2x}(e^{2x} - 1)}{(e^{2x} + 1)^2} = \frac{4e^{2x}}{(e^{2x} + 1)^2} \\
g(x) &= -log(x) + x \\
g'(x) &= -\frac{1}{x} + 1 \\
g''(x) &= \frac{1}{x^2}.
\end{aligned}
$$

De la ecuación, $f'(x) = 0$ sacamos que $x = 0$ es el único punto crítico, lo que se ve también en la figura 12. De la ecuación $g'(x) = 0$ sacamos que $x = 1$ es el único punto crítico, lo que se ve en la figura 12. Para saber si es mínimo o máximo, evaluamos en la segunda derivada. Entonces tenemos que $f''(0) = 1$ y que $g''(1) = 1$, por lo que sabemos que los puntos son mínimos locales. Para garantizar que son mínimos globales, basta ver que $g''(x) \geq 0 \forall x \in \mathbb{R}$ y que $f''(x) \geq 0 \forall x \in \mathbb{R}$, la primera viene del hecho que $x^2 \geq 0 \forall x \in \mathbb{R}$ y la segunda de que $e^x \geq 0 \forall x \in \mathbb{R}$. Además, cabe destacar que la primera nos basta que sea para $(0, \infty)$, ya que este es el dominio de $g(x)$.

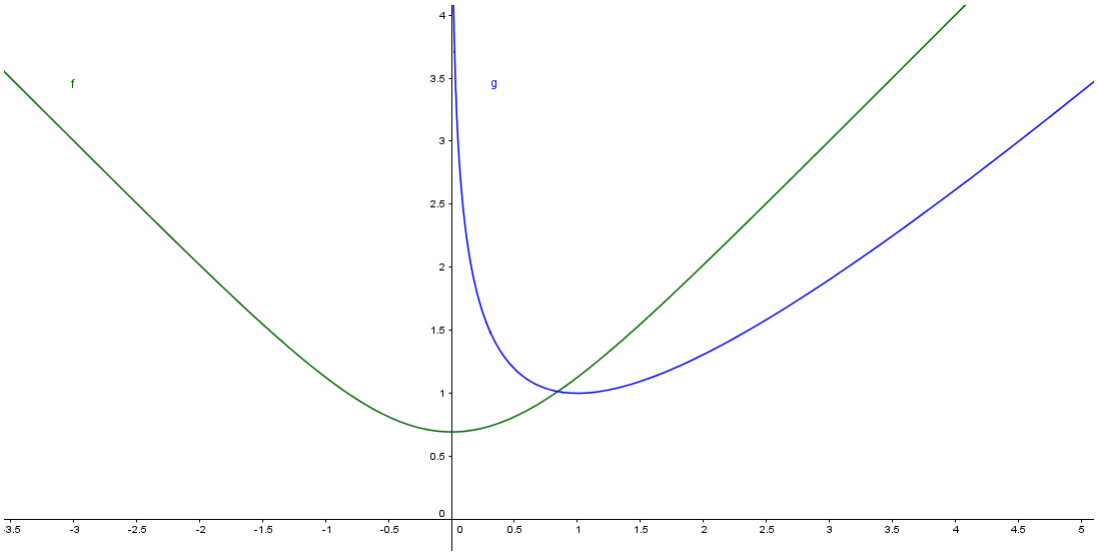


Figura 12: Gráfica de $f(x)$ en verde y $g(x)$ en azul.

ii) Primero para $f(x)$, al iterar desde el punto 1.2 tenemos que:

$$
\begin{aligned}
x^{(1)} &= 1{,}2, & d^{(1)} &= -2{,}31629 \\
x^{(2)} &= -1{,}11629, & d^{(2)} &= 1{,}90101 \\
x^{(3)} &= 0{,}784723, & d^{(3)} &= -0{,}821252 \\
x^{(4)} &= -0{,}0365288, & d^{(4)} &= 0{,}018305 \\
x^{(5)} &= -0{,}0182238, & d^{(5)} &= 0{,}00911693
\end{aligned}
$$

Luego, al iterar sobre el punto 1.3, nos queda que:

$$
\begin{aligned}
x^{(1)} &= 1{,}3, & d^{(1)} &= -2{,}9165 \\
x^{(2)} &= -1{,}6165, & d^{(2)} &= 5{,}86704 \\
x^{(3)} &= 4{,}25054, & d^{(3)} &= -1229{,}52 \\
x^{(4)} &= -1225{,}26, & d^{(4)} &= NaN
\end{aligned}
$$

El problema se da que en esta iteración ya llegamos a valores de e^1000, lo cual no se puede calcular en un pc.

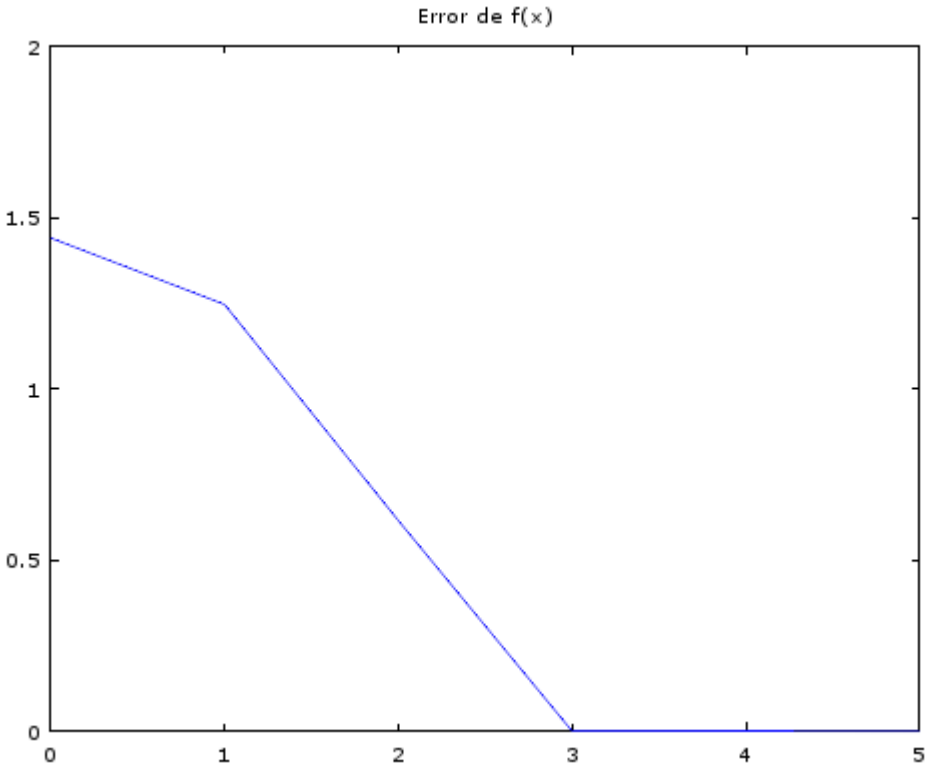


Figura 13: Error del método de Newton en $f(x)$ partiendo del 1.2.

Al resolver $g(x)$ sobre el punto inicial de 1.1 tenemos que:

$$\begin{array}{cc} x^{(1)} = 1{,}1, & d^{(1)} = 2{,}31 \\ x^{(2)} = 3{,}41, & d^{(2)} = 15{,}0381 \\ x^{(3)} = 18{,}4481, & d^{(3)} = 358{,}78 \\ x^{(4)} = 377{,}229, & d^{(4)} = 142679 \\ x^{(5)} = 143056, & d^{(5)} = 2{,}04651e+010 \end{array}$$

Lo que se ve que también diverge.

iii) El gráfico del error es interesante en el caso de $f(x)$ partiendo en 1.2, ya que los otros divergen y no se ve nada. Aún así, no se puede decir mucho del caso de interés, ya que solo contamos con 4 o 5 puntos. Por lo que solo podemos ver que el error disminuye. El gráfico mencionado se ve en la figura 13.

Problema 3. (15 puntos)

a) (6 puntos) Considere el problema

$$\begin{array}{ll} \text{mín} & (x-12)^2 + (y+6)^2 + 10 \\ \text{s.a.} & 2x^2 + 6x + 2y^2 - 9y \leq 13 \\ & (x-9)^2 + y^2 \leq 64 \\ & 8x + 4y = 20, \end{array}$$

se le pide que responda lo siguiente.

i) (2 puntos) Escriba las condiciones de KKT.

ii) (2 punto) Verifique si el punto $(\overline{x}, \overline{y}) = (2, 1)$ cumple las condiciones de KKT.

iii) (2 puntos) ¿El punto $(\overline{x}, \overline{y})$ es óptimo?

b) (3 puntos) Dé ejemplos donde se cumpla que:

i) (1 punto) Tenga un óptimo global no regular.

ii) (1 punto) Existe un punto que cumple las condiciones de KKT, pero no es regular.

iii) (1 punto) Existe exactamente un mínimo local y no existen mínimos globales.

c) (6 puntos) Usted es contratado por una empresa que produce y distribuye bebidas llamada SSU. Se venden n tipos de bebidas distintas, las cuales son producidas en las fábricas que se encuentran en el conjunto I y distribuidas a lo largo de m ciudades. Para cada planta i en I, usted sabe que producir una unidad del producto k cuesta v_{ik} y transportarlo hacia la ciudad j cuesta c_{ijk}. Por temas de tamaño y tecnologías, las fábricas tienen una cota máxima de producción para el producto k de M_{ik}. Además, cada planta tiene una mano de obra de Q_i para producir productos, los cuales deben ser repartidos para saber cuántos producir de cada tipo. Cada producto k gasta q_{ik} horas hombre por cada unidad producida en la planta i.

Por ultimo, debes poder cubrir la demanda de cada ciudad, la cual está dada por d_{jk} que es la demanda del producto k en la ciudad j.

i) (4 puntos) Modele el problema con el objetivo de cumplir la demanda, no romper ninguna restricción de producción y minimizar los costos.

ii) (2 puntos) Cómo cambia el modelo si es que la empresa te dice que tomes en cuenta que poseen p camiones y deben ser asignados al principio de cada día para lograr despachar los pedidos (asuma que viajan por la noche de cualquier ciudad a cualquier ciudad). La información que sabes de los camiones es que el camión h pueden llevar r_h bebidas distintas, independiente del tipo que sean. Se le pide agregar algo al modelo que le permita distribuir los camiones cada día para lograr sacar los productos generados ese mismo día y que puedan ser despachados. Asuma que si un vehículo sale con pedidos, los podrá entregar todos independientemente de a qué ciudad(es) deba ir y luego irá al lugar al que lo envíe.

Solución Problema 3.

a) *i*) Para las condiciones de KKT, primero necesitamos la función Lagrangeano,

$$\mathcal{L}(x, y, \mu_1, \mu_2, \mu_3) = (x-12)^2+(y+6)^2+10+\mu_1(2x^2+6x+2y^2-9y-13)+\mu_2((x-9)^2+y^2-64)+\mu_3(8x+4y-20).$$

Luego, las condiciones de KKT son:

$$
\begin{array}{lccccc}
i) & \frac{\partial \mathcal{L}}{\partial x} & = & 2x - 24y + \mu_1(4x+6) + \mu_2(2x-18) + 8\mu_3 & = & 0 \\
ii) & \frac{\partial \mathcal{L}}{\partial y} & = & 2y + 12 + \mu_1(4y-9) + \mu_2(2y) + 4\mu_3 & = & 0 \\
iii) & \frac{\partial \mathcal{L}}{\partial \mu_1} & = & 2x^2 + 6x + 2y^2 - 9y - 13 & \leq & 0 \\
iv) & \frac{\partial \mathcal{L}}{\partial \mu_2} & = & (x-9)^2 + y^2 - 64 & \leq & 0 \\
v) & \frac{\partial \mathcal{L}}{\partial \mu_3} & = & 8x + 4y - 20 & = & 0 \\
vi) & x\frac{\partial \mathcal{L}}{\partial x} & = & x(2x - 24y + \mu_1(4x+6) + \mu_2(2x-18) + 8\mu_3) & = & 0 \\
vii) & y\frac{\partial \mathcal{L}}{\partial y} & = & y(2y + 12 + \mu_1(4y-9) + \mu_2(2y) + 4\mu_3) & = & 0 \\
viii) & \mu_1\frac{\partial \mathcal{L}}{\partial \mu_1} & = & \mu_1(2x^2 + 6x + 2y^2 - 9y - 13) & = & 0 \\
ix) & \mu_2\frac{\partial \mathcal{L}}{\partial \mu_2} & = & \mu_2((x-9)^2 + y^2 - 64) & = & 0 \\
x) & & & \mu_1 & \geq & 0 \\
xi) & & & \mu_2 & \geq & 0.
\end{array}
$$

ii) Podemos verificar evaluando el punto $(\overline{x}, \overline{y}) = (2, 1)$ en cada una de las condiciones de KKT, lo que nos deja:

$$
\begin{array}{lccc}
i) & -20 + 14\mu_1 - 14\mu_2 + 8\mu_3 & = & 0 \\
ii) & 14 - 5\mu_1 + 2\mu_2 + 4\mu_3 & = & 0 \\
iii) & 0 & \leq & 0 \\
iv) & -14 & \leq & 0 \\
v) & 0 & = & 0 \\
vi) & 2(-20 + 14\mu_1 - 14\mu_2 + 8\mu_3) & = & 0 \\
vii) & 1(14 - 5\mu_1 + 2\mu_2 + 4\mu_3) & = & 0 \\
viii) & \mu_1(0) & = & 0 \\
ix) & \mu_2(-14) & = & 0 \\
x) & \mu_1 & \geq & 0 \\
xi) & \mu_2 & \geq & 0.
\end{array}
$$

Es fácil ver que $\mu_2 = 0$ por $ix)$. Además, podemos verificar que con $\mu_1 = 4$ y $\mu_3 = -1$ el punto cumple las condiciones de KKT.

iii) Como la función objetivo y las primeras restricciones son convexas y la tercera es lineal, sabemos que el problema es convexo, lo que asegura que es una condición suficiente para asegurar que el punto es un óptimo global.

b) *i)* El problema

$$
\begin{array}{ll}
\text{mín} & x \\
\text{s.a} & (x-1)^2 + (y-1)^2 <= 1 \\
& (x+1)^2 + (y-1)^2 <= 1,
\end{array}
$$

cumple que el único punto factible es (1,0), cumple KKT y no es regular.

ii) El problema

$$
\begin{array}{ll}
\text{mín} & -x_2 \\
\text{s.a.} & x_1 - 2(1-x_2)^3 \leq 0 \\
& -x_1 - 2(1-x_2)^3 \leq 0 \\
& x_1, x_2 \geq 0,
\end{array}
$$

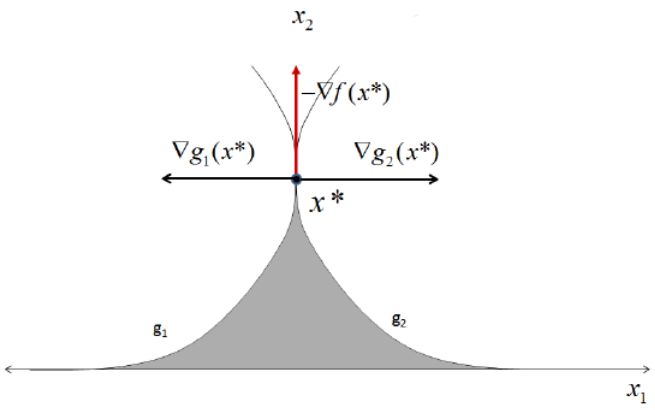


Figura 14: Geometría de un problema que no cumple KKT.

tiene la geometría que se muestra en la figura 14, en donde se ve que no cumple las condiciones de KKT, ya que el gradiente de la función objetivo no se puede escribir como una combinación de los gradientes de las restricciones activas. Este ejemplo se encuentra en detalle en las notas *Constrained Optimality Conditions* del profesor Jorge Vera, en la figura, f es la función objetivo y g_i es la i-ésima restricción.

iii) Considere la función

$$f(x) = \begin{cases} -cos(x), & x \in [\frac{-3\pi}{2}, \frac{3\pi}{2}] \\ x + \frac{3\pi}{2}, & x \in (-\infty, \frac{-3\pi}{2}) \\ -x + \frac{3\pi}{2}, & x \in (\frac{-3\pi}{2}, \infty). \end{cases} .$$

Al minimizar esta función, tendrás que el punto $x = 0$ es el único óptimo local, pero que no tiene óptimo global.

c) **Variables de decisión:**

$$\begin{aligned} x_{ijk} = &\text{ Cantidad de bebida } k \text{ a trasportar de la planta } i \text{ a la ciudad } j. \\ z_{ik} = &\text{ Cantidad de bebida } k \text{ producida en la planta } i. \end{aligned}$$

Función Objetivo:

$$\text{mín} \sum_{i \in I} \sum_{j=1}^{m} \sum_{k=1}^{n} c_{ijk} x_{ijk} \sum_{k=1}^{n} \sum_{i \in I} v_{ik} z_{ik}$$

Restricciones:

- ◇ Lo que se produce se envía a alguna ciudad.

$$z_{ik} = \sum_{j=1}^{m} x_{ijk} \quad \forall i \in I, k = 1, .., n.$$

- ◇ No excedo la mano de obra.

$$\sum_{k=1}^{n} q_{ik} z_{ik} \leq Q_i \quad \forall i \in I.$$

- ◇ Cubro la demanda.

$$\sum_{i \in I} x_{ijk} = d_{jk} \quad \forall j = 1, .., m, k = 1, .., n.$$

- ◇ No produzco mas que la cota dada por tecnología en cada fabrica.

$$z_{ik} \leq M_{ik} \quad \forall i \in I, k = 1, .., n.$$

◇ Naturaleza de las variables.

$$x_{ijk} \geq 0, z_{ik} \geq 0 \quad \forall i \in I, j = 1, .., m, k = 1, .., n.$$

b) Para modelar esto debemos agregar variables:

$$y_{ijkh} = \text{ Cantidad de bebida } k \text{ trasporta el camión } h \text{ de la planta } i \text{ a la ciudad } j.$$

$$w_{ih} = \begin{cases} 1, \text{ si el camión } h \text{ parte al principio del día en la planta } i. \\ 0, \text{ en otro caso.} \end{cases}$$

Las restricciones que agregaremos con estas variables son:

◇ Transporto todo lo que se produce en la planta i y se envía a la ciudad j.

$$\sum_{h=1}^{p} \sum_{k=1}^{n} y_{ijkh} = \sum_{k=1}^{n} x_{ijk} \quad \forall i \in I, j = 1, .., m.$$

Ojo que esto asegura que todo lo que se produce en la planta i sale, ya que todo se envía a ciudades j's.

◇ Definir que el camión sale de una ciudad solamente y no se desdobla.

$$\sum_{i \in I} w_{ih} = 1 \forall h = 1, .., p$$

◇ Si el camión no parte en i, entonces no puede sacar bebidas de ahí.

$$\sum_{j=1}^{m} \sum_{k=1}^{n} y_{ijkh} \leq L w_{ih} \forall i \in I, h = 1, .., p.$$

Donde L es una constante suficientemente grande.

Problema 4. (15 puntos)

a) (8 puntos) Usted quiere hacer una fiesta para despedir un amigo llamado Juan y se le ha designado la tarea de ver a cuántas personas invitar y a quiénes. Usted sabe que los amigos de Juan están en el conjunto J (usted esta dentro de J), que los amigos tuyos son las personas del conjunto Y y que la fiesta la puede hacer en la casa de alguien que sea amigo de ambos, ya que usted le pedirá que preste la casa (tu casa también esta disponible para hacer la fiesta). Usted es el encargado de decidir en qué casa se hará la fiesta, y para eso definió c_i como la capacidad de personas que caben en la casa de la persona i.

Un amigo que te está ayudando, te dice que tengas en cuenta que si le pides a la persona i que preste la casa para la fiesta, entonces la persona debe tener incentivos para eso. Esto se traduce en que los que te prestarán la casa serán los que son amigos tuyos y amigos de Juan. Además, si le pides la casa a la persona i, ésta te exigirá que invites a sus amigos, los cuales están en el conjunto I_i.

Investigando sobre Juan, lograste descubrir que él siente aprecio distinto por sus amigos y eso se puede cuantificar en que él obtiene b_i de felicidad si es que su amigo i va a la fiesta y v_i si es que un desconocido va a la fiesta.

Por otro lado, sabes que algunas personas tienen problemas entre ellas y lo puedes resumir en que hay cierto tipo de riña entre las personas del grupo P y el grupo L y esto lo descubriste ya que te dijeron los del grupo P que si invitas más personas del grupo L que las de su grupo, entonces que no te molestes en invitarlos. Además, las personas del conjunto D te dicen que ellas solo van a fiestas que consideran "masivas", lo que se traduce en que para cada i en D, solo irán a la fiesta si es que se invita al menos a d_i personas.

Se le pide modelar un problema lineal entero mixto, con el que puedas elegir a quiénes invitar a la fiesta y dónde hacerla con el objetivo de maximizar la felicidad de Juan.

b) (3 puntos) Considere $S \subset \mathbb{R}$ es convexo y $f : S \to \mathbb{R}$. Demuestre f es convexa si y solo si se cumple que

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \quad \forall x, y \in S.$$

c) (4 puntos) Demuestre que si x_1 y x_2 son dos soluciones consecutivas del algoritmo del gradiente, entonces se cumple que:

$$\nabla f(x_1)^T \nabla f(x_2) = 0,$$

con f la función objetivo.

Solución Problema 4.

a) Primero, haremos unos arreglos en los parámetros y conjuntos para que sea mas fácil trabajar con ellos. Primero, me agregare a mi a Y, en el caso de que esté incluido (no hago nada), sino, reescribo Y agregándome a mí. Por otro lado, generare el conjunto T, que contendrá a todas las personas, tanto en Y, J y los I_i para todos los i en $J \cap Y$. Si quiero que Juan pueda prestar la casa, es cosa de asegurarme que este en J y en Y, pero eso no afecta el modelo.

Además, se asume que los que invito a la fiesta irán, a menos que ya sepa a priori que no irán (por ejemplo, no asumiré que los de P irán si es que invito a más del grupo D).

Variables:

$$y_i = \begin{cases} 1 \text{ , si se ocupa la casa de } i \in Y \cap Y \\ 0 \text{ , en otro caso.} \end{cases}$$

$$x_i = \begin{cases} 1 \text{ , si se invita a la persona } j \in T \\ 0 \text{ , en otro caso.} \end{cases}$$

$$z = \begin{cases} 1 \text{ , si se privilegia al grupo } P \\ 0 \text{ , en otro caso.} \end{cases}$$

Función objetivo:

$$\text{máx} \sum_{i \in J} b_i x_i + \sum_{i \in T \setminus J} v_i x_i$$

Restricciones:

- ◇ Tengo que invitar a los amigos de la persona que preste la casa.

$$\sum_{j \in I_i} x_j \geq |I_i| y_i \quad \forall i \in Y \cap J.$$

- ◇ Solo una persona prestará la casa.

$$\sum_{i \in Y \cap J} y_i = 1.$$

- ◇ No puedo invitar a mas gente que la capacidad de la casa.

$$\sum_{j \in T} x_j \leq \sum_{i \in J \cap Y} y_i c_i.$$

- ◇ Si no privilegio a P, mejor no los invito para no calcular mal el espacio y el beneficio.

$$\sum_{i \in P} x_i \leq |P| z.$$

◇ Darle sentido a la variable z.

$$\sum_{i \in P} x_i \geq z \sum_{i \in L} x_i.$$

Ojo que la parte derecha es multiplicación de variables binarias, lo que sabemos que es linealizable al agregar variables binarias auxiliares.

◇ Invitar a algún tipo en D solo si la fiesta es masiva, así no arruino mis cálculos de beneficio y capacidad de la casa.

$$\sum_{i \in T} x_i \geq d_j x_j \quad \forall j \in D.$$

◇ Naturaleza de variables.

$$\begin{aligned} x_i &\in \{0,1\} \forall i \in T \\ y_j &\in \{0,1\} \forall j \in J \cap Y \\ z &\in \{0,1\}. \end{aligned}$$

b) Como el $S \subset \mathbb{R}$ y f va de $\mathbb{R}$ a $\mathbb{R}$, entonces la demostración se reduce a demostrar con la condición:

$$f(y) \geq f(x) + f'(x)^T (y - x), \quad \forall x, y \in S. \tag{1}$$

Deberíamos asumir que f es diferenciable, en este caso:

⇒ Como f es convexa y S es convexa, tenemos que si $x, y \in S$, entonces $ty + (1-t)x \in S$ para $0 \leq t \leq 1$. Además, tenemos que $ty + (1-t)x = x + t(y - x)$, por lo que por convexidad:

$$\begin{aligned} & f(x + t(y-x)) \leq (1-t)f(x) + tf(y) \\ \Rightarrow \qquad & f(y) \geq f(x) + \frac{f(x + t(y-x)) - f(x)}{t}, \end{aligned}$$

y tomando limite con $t \to 0$, tenemos lo pedido.

⇐ Para este lado, asumimos que que se cumple (1) para todos los $x \in S$. Sea $x \neq y$ en S y $0 \leq \theta \leq 1$ y sea $z = \theta x + (1-\theta)y$. Aplicando (1) dos veces, tenemos:

$$\begin{aligned} f(x) &\geq f(z) + f'(z)(x - z) \\ f(y) &\geq f(z) + f'(z)(y - z) \end{aligned}$$

multiplicando la primera ecuación por θ y la segunda por $(1-\theta)$ y sumandolas nos da que:

$$\theta f(x) + (1+\theta) f(y) \geq f(z)$$

que es lo pedido.

c) Sabemos que $x_2 = x_1 - \nabla f(x_1)\lambda_1$, donde λ_1 cumple que es el óptimo del problema

$$\min_{\lambda} \{f(x_1 - \nabla f(x_1)\lambda)\},$$

por lo que sabemos que debe cumplir que al derivar por λ la función objetivo, debe dar 0. Además, sabemos que por regla de la cadena,

$$\frac{\partial f(x_1 - \nabla f(x_1)\lambda)}{\partial \lambda} = \nabla f(x_1 - \nabla f(x_1)\lambda)(-\nabla f(x_1).$$

Usando esos dos argumentos, tenemos lo pedido.