Optimización

Condiciones de Karush-Kuhn-Tucker

Dr. E Uresti

ITESM

Las condiciones necesarias que deben satisfacer los óptimos de problemas de optimización no lineal con restricciones de desigualdad fueron publicadas por primera vez (1939) en la tesis de Maestría de William Karush (1917-1997) (en aquél entonces estudiante de matemáticas de la Universidad de Chicago) , aunque fueron renombradas tras un artículo en una conferencia de Harold W. Kuhn y Albert W. Tucker en 1951. Las condiciones de Karush-Kuhn-Tucker (KKT) son una generalización del método de los multiplicadores de Lagrange para restricciones de desigualdad.

Considere el problema de optimización

$$\text{Min } f(x_1, x_2, \ldots, x_n)$$

sujeto a

$$\begin{array}{rcl} g_1(x_1, x_2, \ldots, x_n) & \leq & 0 \\ g_2(x_1, x_2, \ldots, x_n) & \leq & 0 \\ & \vdots & \\ g_m(x_1, x_2, \ldots, x_n) & \leq & 0 \end{array} \tag{0.1}$$

El método de solución procede de la siguiente manera. Cambiemos cada restricción de desigualdad $g_i \leq 0$ a una restricción de igualdad introduciendo una variable s_i de la siguiente manera:

$$g_i \leq 0 \rightarrow g_i + s_i^2 = 0$$

De acuerdo a la técnica de los multiplicadores de Lagrange se construye la función:

$$\boxed{F(\mathbf{x}, \lambda, \mathbf{s}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \cdot (g_i + s_i^2)}$$

Los puntos que minimizan a f sujeta a las restricciones $g_i \leq 0$ $(1 \leq i \leq m)$ están dentro de los puntos críticos de F:

- Que hacen cero las parciales con respecto a las variables x_j $(j = 1, \ldots, n)$:

$$\frac{\partial F}{\partial x_j} = \frac{\partial f}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0$$

- Que hacen cero las parciales con respecto a las variables λ_i $(i = 1, \ldots, m)$:

$$\frac{\partial F}{\partial \lambda_i} = g_i + s_i^2 = 0 \leftrightarrow g_i \leq 0$$

- Que hacen cero las parciales con respecto a las variables s_i $(i = 1, \ldots, m)$:

$$\frac{\partial F}{\partial s_i} = 2\,\lambda_i\, s_i = 0 \leftrightarrow \lambda_i\, s_i = 0 \leftrightarrow \lambda_i\, g_i = 0$$

Teorema

Suponga una formulación para el problema anterior de minimización. Si $\mathbf{x}_0 = (a_1, a_2, \ldots, a_n)$ es un óptimo, entonces deben existir números reales llamados multiplicadores λ_1, λ_2,...,λ_m no negativos tales que $(a_1, a_2, \ldots, a_n, \lambda_1, \ldots, \lambda_m)$ es un punto crítico para F. Es decir que se cumple:

$$
\begin{array}{c}
\text{Bloque I} \\
+\frac{\partial f(\mathbf{x}_o)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\mathbf{x}_o)}{\partial x_j} = 0 \quad j = 1, 2 \ldots, n \\
\text{Bloque II: Condición de Holgura Complementaria} \\
\lambda_i \, g_i(\mathbf{x}_o) = 0 \quad i = 1, 2, \ldots, m \\
\text{Bloque III} \\
g_i \leq 0 \quad i = 1, 2, \ldots, m
\end{array}
\tag{0.2}
$$

Observe que los valores de s_i se obtienen de la relación $g_i + s_i^2 = 0$ y de que $g_i \leq 0$.

Si ahora el problema es de maximización:

$$\text{Max } f(x_1, x_2, \ldots, x_n)$$

sujeto a

$$\begin{aligned} g_1(x_1, x_2, \ldots, x_n) &\leq 0 \\ g_2(x_1, x_2, \ldots, x_n) &\leq 0 \\ &\vdots \\ g_m(x_1, x_2, \ldots, x_n) &\leq 0 \end{aligned} \tag{0.3}$$

Para su solución lo cambiamos a un problema de minimización para $-f(\mathbf{x})$. En este caso la función F queda en la forma:

$$F(\mathbf{x}, \lambda, \mathbf{s}) = -f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \cdot (g_i + s_i^2)$$

Teorema

Suponga una formulación para el problema anterior en el caso de maximización. Si $\mathbf{x}_0 = (a_1, a_2, \ldots, a_n)$ es un óptimo, entonces deben existir números reales llamados multiplicadores λ_1, λ_2,$\ldots$,λ_m no negativos tales que $(a_1, a_2, \ldots, a_n, \lambda_1, \ldots, \lambda_m)$ es un punto crítico para F. Es decir, que se cumple:

$$
\begin{array}{rcll}
 & \text{Bloque I} & & \\
-\frac{\partial f(\mathbf{x}_o)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\mathbf{x}_o)}{\partial x_j} & = & 0 & j = 1, 2 \ldots, n \\
 & \text{Bloque II} & & \\
\lambda_i \, g_i(\mathbf{x}_o) & = & 0 & i = 1, 2, \ldots, m \\
 & \text{Bloque III} & & \\
g_i & \leq & 0 & i = 1, 2, \ldots, m
\end{array}
\tag{0.4}
$$

Uso de las Condiciones KKT

La forma de operar las condiciones de KKT será la siguiente: Como lo que buscamos es el punto $\mathbf{x}_o$ y de inicio se desconoce, entonces las ecuaciones de las condiciones de los bloques I y II se piensan como un sistema de ecuaciones en las variables $x_j{}'s$ y $\lambda_j{}'s$: Se intenta resolver tal sistema de ecuaciones y en caso de encontrarse las soluciones se revisan una a una para ver cual de ella cumple que los $\lambda_j{}'s$ son no negativos y que también se cumplen las restricciones $g_i \leq 0$ en los puntos encontrados. Normalmente se realiza una tabla donde se hace la verificación.

Observe también es posible trabajar el problema de maximización resolviendo el problema de minimización pero conservando aquellos puntos que tengan los valores de los multiplicadores no positivos.

Ejemplo 1

Encuentre los valores mínimo y máximo de la función $f(x_1, x_2) = 3 - x_1 - x_2$ sujeta a las restricciones $0 \leq x_1$, $0 \leq x_2$ y $2\,x_1 + x_2 \leq 2$.

Encuentre los valores mínimo y máximo de la función $f(x_1, x_2) = 3 - x_1 - x_2$ sujeta a las restricciones $0 \leq x_1$, $0 \leq x_2$ y $2\,x_1 + x_2 \leq 2$.

Solución

- Primero cambiemos las restricciones a la forma $g_i \leq 0$:

$$
\begin{aligned}
0 \leq x_1 &\quad\rightarrow\quad g_1 = -x_1 \leq 0 \\
0 \leq x_2 &\quad\rightarrow\quad g_2 = -x_2 \leq 0 \\
x_1 + x_2 \leq 2 &\quad\rightarrow\quad g_3 = 2\,x_1 + x_2 - 2 \leq 0
\end{aligned}
$$

- Resolvamos el problema de minimización primeramente. En este caso las condiciones son:

Bloque I

$$\frac{\partial f(\mathbf{x}_o)}{\partial x_1} + \sum_{i=1}^m \lambda_i \frac{\partial g_i(\mathbf{x}_o)}{\partial x_1} = -1 + 2\,\lambda_1 - \lambda_2 = 0$$

$$\frac{\partial f(\mathbf{x}_o)}{\partial x_2} + \sum_{i=1}^m \lambda_i \frac{\partial g_i(\mathbf{x}_o)}{\partial x_1} = -1 + \lambda_1 - \lambda_3 = 0$$

Bloque II: Condición de Holgura Complementaria

$$\begin{aligned} \lambda_1\, g_1 &= \lambda_1\,(2\,x_1 + x_2 - 2) &= 0 \\ \lambda_2\, g_2 &= -\lambda_2\, x_1 &= 0 \\ \lambda_3\, g_3 &= -\lambda_3\, x_2 &= 0 \end{aligned}$$

x_1	x_2	λ_1	λ_2	λ_3	g_1	g_2	g_3	f
0	0	0	-1	-1	-2	0	0	3
1	0	1/2	0	-1/2	0	-1	0	2
0	2	1	1	0	0	0	-2	1

- Para determinar el máximo las condiciones quedan:

Bloque I

$$
\begin{aligned}
-\frac{\partial f(\mathbf{x}_o)}{\partial x_1} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\mathbf{x}_o)}{\partial x_1} &= 1 + 2\,\lambda_1 - \lambda_2 &= 0 \\
-\frac{\partial f(\mathbf{x}_o)}{\partial x_2} + \sum_{i=1}^{m} \lambda_i \frac{\partial g_i(\mathbf{x}_o)}{\partial x_1} &= 1 + \lambda_1 - \lambda_3 &= 0
\end{aligned}
$$

Bloque II: Condición de Holgura Complementaria

$$
\begin{aligned}
\lambda_1\, g_1 &= \lambda_1\,(2\,x_1 + x_2 - 2) &= 0 \\
\lambda_2\, g_2 &= -\lambda_2\, x_1 &= 0 \\
\lambda_3\, g_3 &= -\lambda_3\, x_2 &= 0
\end{aligned}
$$

x_1	x_2	λ_1	λ_2	λ_3	g_1	g_2	g_3	f
0	0	0	1	1	-2	0	0	3
1	0	-1/2	0	1/2	0	-1	0	2
0	2	-1	-1	0	0	0	-2	1

Observamos que las tablas para minimización y para maximización son idénticas salvo que los valores de los multiplicadores están cambiados de signo. Por tanto, la estrategia conveniente para optimizar una función sujeta a restricciones de desigualdad por el método de las condiciones de KKT será:

1. Plantear el problema **como si se tratará sólo de minimización** y resolver el sistema de ecuaciones correspondientes.
2. Eliminar aquellos puntos encontrados que no satisfacen las restricciones $g_i \leq 0$.
3. Eliminar aquellos puntos que tienen a la vez multiplicadores positivos y negativos.
4. Para minimización: escoger dentro de aquellos puntos que tienen multiplicadores no negativos aquél que tienen la menor evaluación de la función objetivo.
5. Para maximización: escoger dentro de aquellos puntos que tienen multiplicadores no positivos aquél que tienen la mayor evaluación de la función objetivo.

Encuentre los máximos y mínimos **absolutos** de la función:

$$f(x, y) = x^2 + y^2 + y - 1$$

En la región S definida por

$$S = \left\{(x, y) \in \mathbf{R}^2 | x^2 + y^2 \leq 1\right\}$$

Solución

Utilizaremos las condiciones KKT para caracterizar los máximos y los mínimos. Aquí $g = g(x,y) = x^2 + y^2 - 1 \leq 0$. En la figura 1 aparecen los preparativos para la solución del problema, así como sus puntos críticos. El orden de las variables en la matriz es $x - y - t$.

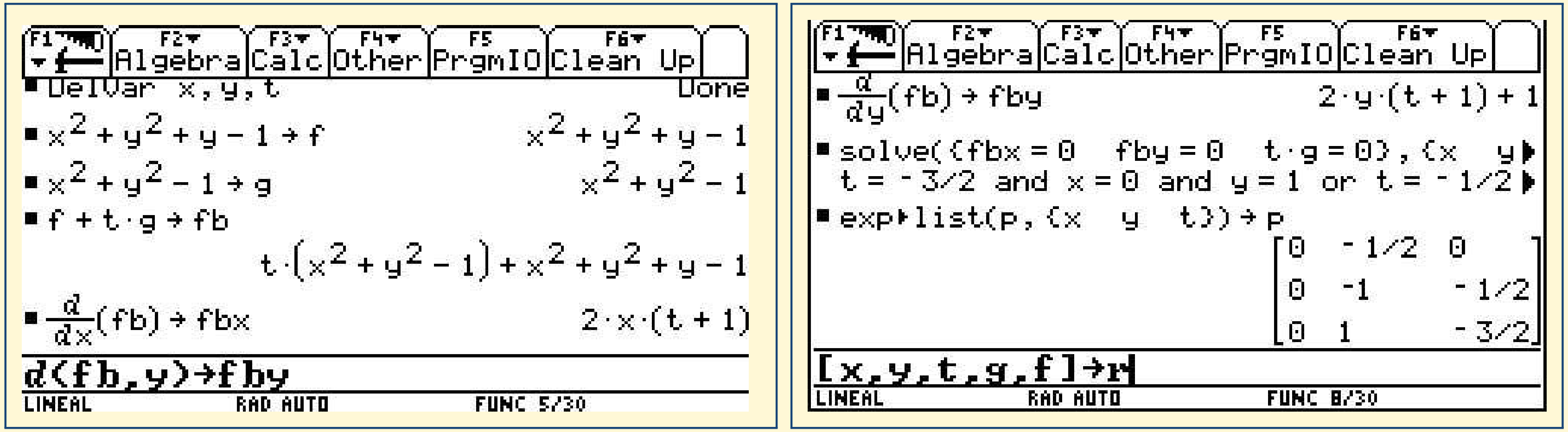


Figura 1: Preparativos y puntos críticos del ejemplo 2

x	y	t	g	f
0	-1/2	0	-3/4	-5/4
0	-1	-1/2	0	-1
0	1	-3/2	0	1

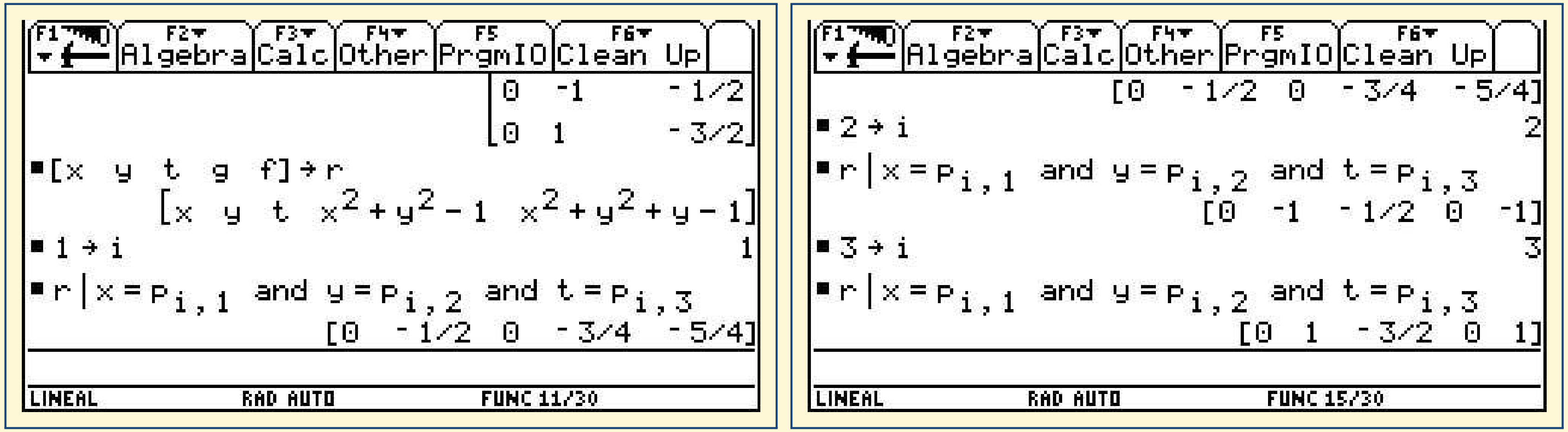


Figura 2: Sustitución de los puntos críticos en $[x, y, t, g, f]$

Por lo tanto, $f(x = 0, y = -1/2) = -5/4$ es el mínimo de la función y $f(x = 0, y = 1) = 1$ es el valor máximo.

Ejemplo 3

Un comerciante puede comprar hasta 17.25 onzas de un producto químico A a 10 dólares cada onza. Se puede convertir una onza del producto químico A en una onza del producto I a un costo de 3 dólares a onza. Asimismo, una onza del químico A se puede convertir en una onza del producto II a un costo de 5 dólares la onza. Si se producen x_1 onzas del producto I se venderá a $30 - x_1$ dólares la onza, mientras que si se producen x_2 onzas del producto II se venderá a $50 - x_2$ dólares la onza. Determine cómo el comerciante puede maximizar sus ganancias.

Ejemplo 3

Un comerciante puede comprar hasta 17.25 onzas de un producto químico A a 10 dólares cada onza. Se puede convertir una onza del producto químico A en una onza del producto I a un costo de 3 dólares a onza. Asimismo, una onza del químico A se puede convertir en una onza del producto II a un costo de 5 dólares la onza. Si se producen x_1 onzas del producto I se venderá a $30 - x_1$ dólares la onza, mientras que si se producen x_2 onzas del producto II se venderá a $50 - x_2$ dólares la onza. Determine cómo el comerciante puede maximizar sus ganancias.

Variables de Decisión

- x_1 = Onzas del producto I producidas
- x_2 = Onzas del producto II producidas

Objetivo

$$\text{Max } z = x_1\,(30 - x_1) + x_2\,(50 - x_2) - 3\,x_1 - 5\,x_2 - 10\,(x_1 + x_2)$$

Restricciones

$$x_1 + x_2 \leq 17.25,\ \ 0 \leq x_1,\ \ 0 \leq x_2$$

Así:

$$
\begin{aligned}
f &= x_1\,(30 - x_1) + x_2\,(50 - x_2) - 3\,x_1 - 5\,x_2 - 10\,(x_1 + x_2) \\
g_1 &= x_1 + x_2 - 17.25 \leq 0 \\
g_2 &= -x_1 \leq 0 \\
g_3 &= -x_2 \leq 0
\end{aligned}
$$

Las condiciones de KKT que debe satisfacer el óptimo son (Observe que se uso el criterio para maximizar con $-f$; por tanto, los multiplicadores no deben ser negativos):

Bloque I

$$
\begin{array}{rcl}
-\frac{\partial f}{\partial x_1} + \sum_{i=1}^{3} \lambda_i \cdot \frac{\partial g_i}{\partial x_1} = -17 + 2\,x_1 + \lambda_1 - \lambda_2 & = & 0 \\
-\frac{\partial f}{\partial x_2} + \sum_{i=1}^{3} \lambda_i \cdot \frac{\partial g_i}{\partial x_1} = -35 + 4\,x_2 + \lambda_1 - \lambda_3 & = & 0
\end{array}
$$

Bloque II

$$
\begin{array}{rcl}
\lambda_1 \cdot (g_1) = \lambda_1\,(x_1 + x_2 - 17.25) & = & 0 \\
\lambda_2 \cdot (g_2) = -\lambda_2\,x_1 & = & 0 \\
\lambda_3 \cdot (g_3) = -\lambda_3\,x_2 & = & 0
\end{array}
$$

Resolviendo el sistema anterior con Maple obtenemos los siguientes puntos. En la tabla se tabula cada una de las restricciones evaluada en el punto correspondiente. Recuerde que las λs deben ser no negativas y las restricciones deben cumplirse ($g_i \leq 0$):

x_1	x_2	λ_1	λ_2	λ_3	$g_1(\mathbf{x})$	$g_2(\mathbf{x})$	$g_2(\mathbf{x})$	$f(\mathbf{x})$
0	0	0	-17.	-35.	-17.25	0	0	0
8.50	0	0	0	-35.	-8.75	-8.50	0	72.25
17.25	0	-17.5	0	-52.5	0	-17.25	0	-4.3125
0	17.5	0	-17.	0	.25	0	-17.5	306.25
8.50	17.5	0	0	0	8.75	-8.50	-17.5	378.5
0	17.25	.500	-16.5	0	0	0	-17.25	306.1875
4.125	13.125	8.75	0	0	0	-4.125	-13.125	340.21875

Por consiguiente, el único punto sobrevieviente es el del renglón 7: $x_1 = 4.125$ y $x_2 = 13.125$ con una evaluación de 340.21875.

Nota

Si se utiliza LINGO para resolver el problema codificándolo como

MAX=x1*(30-x1)+x2*(50-x2)-3*x1-5*x2-10*(x1+x2);

x1+x2$<=$17.25;

se obtiene:

Local optimal solution found.

Objective value: 340.2188

Variable	Value	Reduced Cost
X1	4.12500	0.000000
X2	13.12500	0.0000000

Row	Slack or Surplus	Dual Price
1	340.2188	1.000000
2	0.0000000	-8.75000

Esto coincide con nuestro cálculo.

Hagamos las operaciones utilizando la calculadora TI. En la figura 3 se muestran las pantallas donde inician los preparativos: primeramente se limpian las variables que serán usadas: $x1$, $x2$, $t1$ (en lugar de λ_1), $t2$ (en lugar de λ_2), y $t3$ (en lugar de λ_3). Es conveniente manejar las restricciones en la forma $g_i \leq 0$. Las variables $g1$, $g2$ y $g3$ codificarán los lados izquierdos de las restricciones. Las ecuaciones del bloque I se definirán utilizando variables $e1$ y $e2$ que representan los lados izquierdos de ellas. Así

$$
\begin{aligned}
e1 &= \frac{\partial f}{\partial x1} - \sum_{i=1}^{3} ti \cdot \frac{\partial gi}{\partial x1} \\
e2 &= \frac{\partial f}{\partial x2} - \sum_{i=1}^{3} ti \cdot \frac{\partial gi}{\partial x2}
\end{aligned}
$$

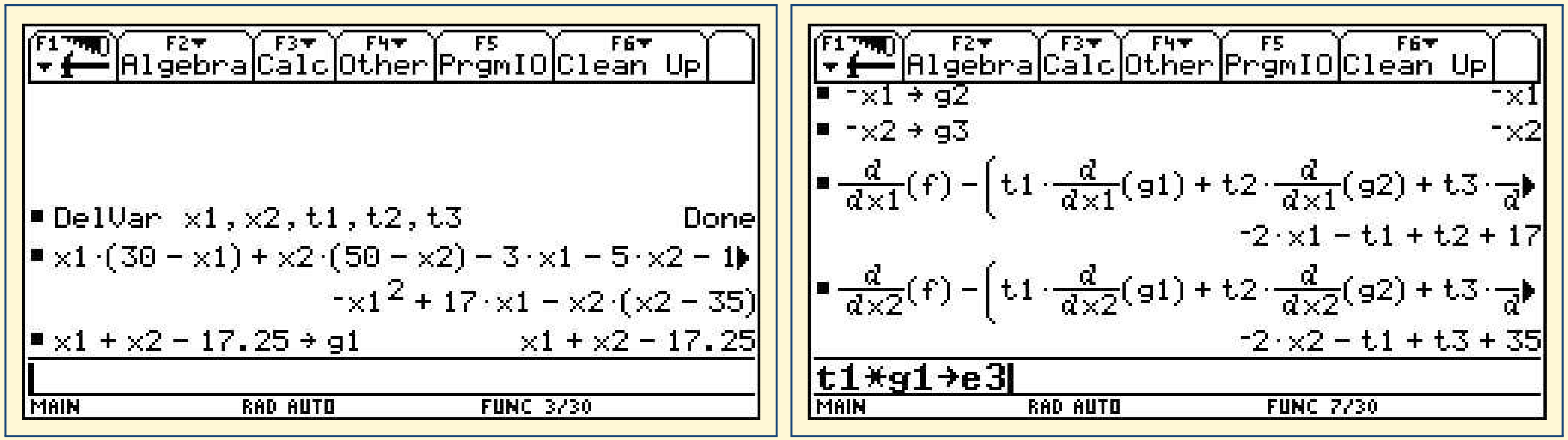


Figura 3: Preparativos en la TI para el ejemplo 3

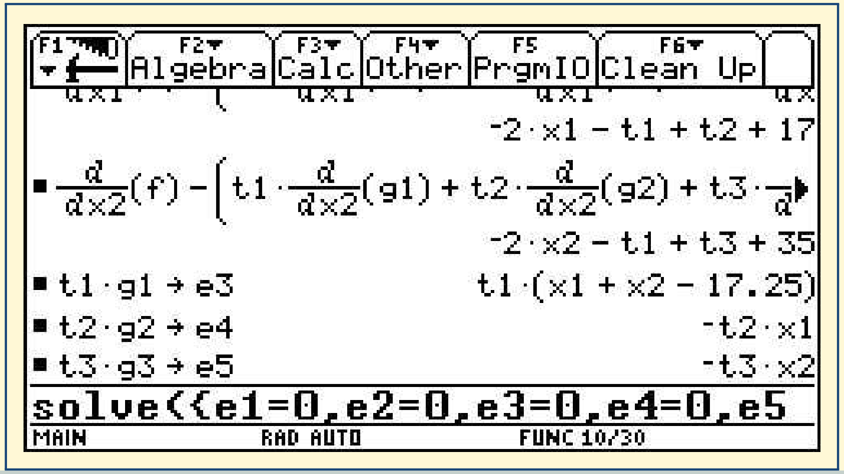
En la figura 4 se muestran las ecuaciones del bloque II y la solución del sistema para los puntos críticos así como su conversión a matriz.

$$
\begin{aligned}
e3 &= t1 \cdot g1\\
e4 &= t2 \cdot g2\\
e5 &= t3 \cdot g3
\end{aligned}
$$

Sistema

$$
\begin{aligned}
e1 &= 0\\
e2 &= 0\\
e3 &= 0\\
e4 &= 0\\
e5 &= 0
\end{aligned}
$$

Para $x1, x2, t1, t2, t3$



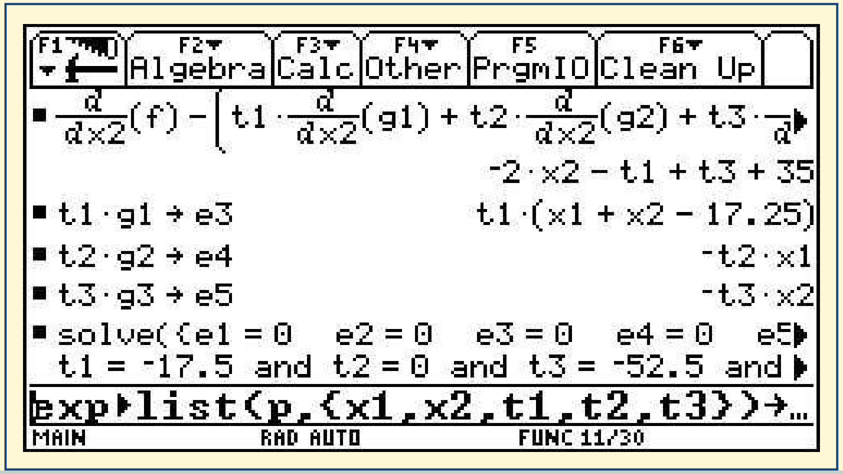


Figura 4: Formación del sistema para los puntos críticos del ejemplo 3

En la figura 5 se muestran las raíces del sistema que define los puntos críticos. Observe que estas 7 raíces coinciden con los resultados de Maple. Recuerde que en la primer columna aparece el valor de $x1$, en la segunda el de $x2$, en la tercera el de $t1$, en la cuarta el de $t2$ y en la quinta el de $t3$. Como los valores de ti esperados deben ser positivos esto descarta todos excepto los correspondientes a los renglones 1 y 3: $P(x1 = 4.125, x2 = 13.125, t1 = 8.75, t2 = 0, t3 = 0)$ y $Q(x1 = 8.5, x2 = 17.5, t1 = 0, t2 = 0, t3 = 0)$

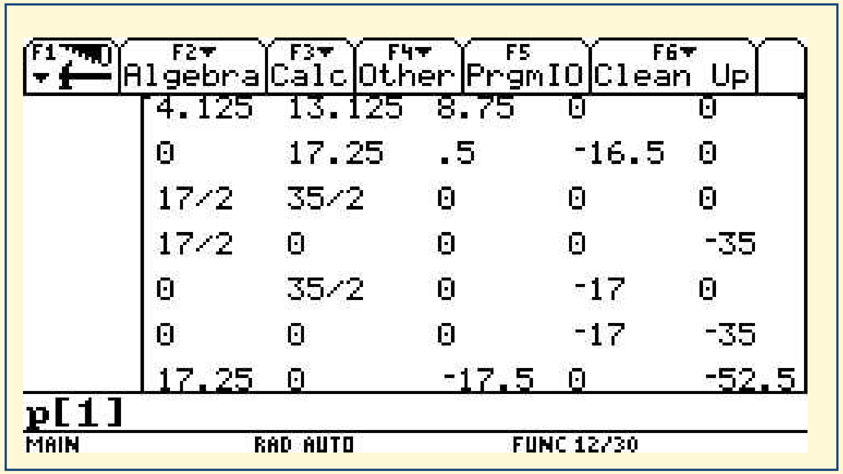


Figura 5: Puntos críticos del ejemplo 3

Recuerde que algunos de estos puntos críticos pueden no cumplir las restricciones y deben pasarse por la verificación. En la figura 6 se muestran los vectores $[g1, g2, g3]$ resultantes de sustituir en las restricciones los puntos. Recuerde que las restricciones son del tipo $gi \leq 0$. Obervamos que el punto Q se descarta pues $g1(Q) = 8.75 > 0$ y que el punto P las cumple. Por tanto, el único punto máximo de f sujeto a las restricciones es $P(x1 = 4.125, x2 = 13.125, t1 = 8.75, t2 = 0, t3 = 0)$ y que tiene una evaluación de 340.219 □

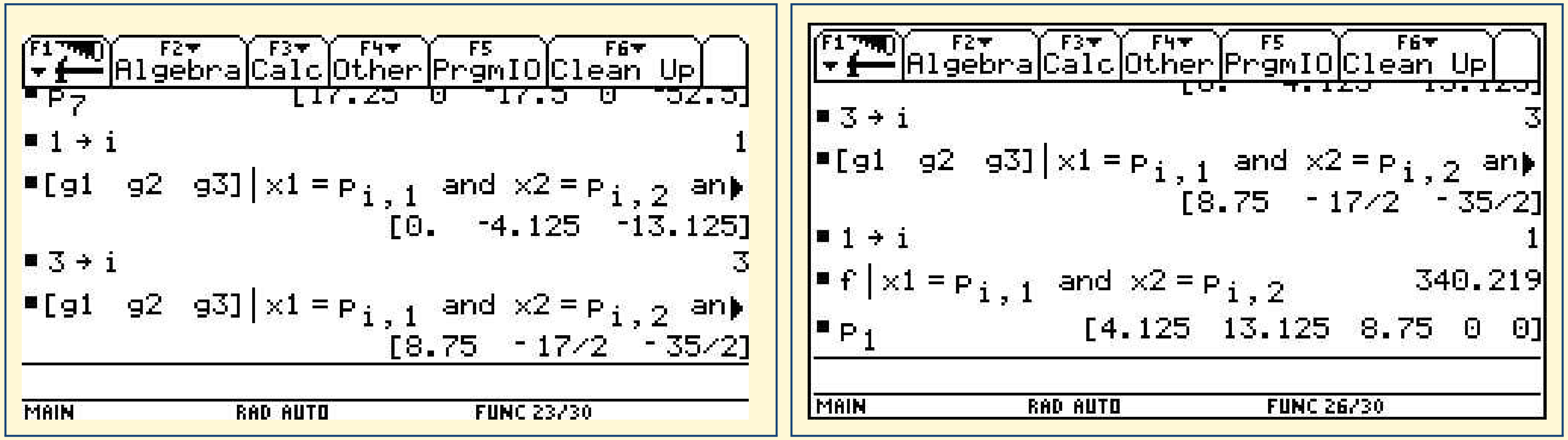


Figura 6: Verificación de restricciones para el ejemplo 3

Los siguientes serán los ejercicios de tarea para este tema.

Ejercicio 1

Utilizando las condiciones de KKT resuelva el problema:

$$\text{Max } z = x_1 - x_2$$

sujeto a la condición:

$$g_1(x_1, x_2) = {x_1}^2 + {x_2}^2 - 1 \leq 0$$

Indique en orden los valores de x_1, x_2 y de z.

Ejercicio 2

Utilice las condiciones de KKT para encontrar la solución óptima del siguiente problema:

$$\text{Min } z = (x_1 - 3)^2 + (x_2 - 5)^2$$

sujeto a:

$$\begin{aligned} g(x_1, x_2) &= x_1 + x_2 - 7 \leq 0 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

Indique en orden los valores de x_1, x_2 y de z.

Ejercicio 3

Utilice las condiciones de KKT para encontrar la solución óptima del siguiente problema:

$$\text{Max } z = (x_1 - 1)^2 + (x_2 - 2)^2$$

sujeto a:

$$\begin{aligned} g_1(x_1, x_2) &= -x_1 + x_2 - 1 = 0 \\ g_2(x_1, x_2) &= x_1 + x_2 - 2 \leq 0 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

Indique en orden los valores de x_1, x_2 y de z.

Sugerencia: Codifique la restricción $g_1 = 0$ mediante las dos restricciones $g_1 \leq 0$ y $g_1 \geq 0$ ($-g_1 \leq 0$).

Ejercicio 4

Una compañía de energía eléctrica enfrenta demandas de energía durante los tiempos de carga máxima y no máxima. Si cobra un precio de p_1 dólares el kilowatt-hora durante el tiempo de carga máxima, entonces los clientes pedirán $60 - 0.5\,p_1$ kwh de energía. Si se cobra un precio de p_2 dólares el kilowatt-hora, entonces los clientes pedirán $40 - p_2$ kwh. La compañía tiene que tener la suficiente capacidad para satisfacer la demanda durante los dos tiempos. A la compañía le cuesta 10 dólares al día mantener cada kilowatt-hora de capacidad. Determine cómo la compañía puede maximizar los ingresos diarios menos los costos de operación. Indique

- la capacidad de la compañía en kwh,
- el precio en dólares durante el tiempo de demanda máxima, y
- el precio en dólares fuera de demanda máxima.

Ejercicio 5

Se disponen semanalmente de un total de 160 horas de mano de obra a 15 dólares la hora. Se puede conseguir mano de obra adicional a 25 dólares la hora. Se puede obtener capital en cantidades ilimitadas a un costo de 5 dólares la unidad de capital. Si se disponen de K unidades de capital y de L horas de mano de obra, entonces se pueden producir $L^{1/2} K^{1/3}$ máquinas. Se vende cada máquina a 270 dólares. ¿Cómo puede la empresa maximizar sus ganancias? Indique

- el total de horas de mano de obra a utilizar,
- el total de unidades de capital, y
- el total de máquinas a producir.