Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Matemática
MA3701-2 Optimización

Pauta Control 2

Profesores: Vicente Acuña, Sebastián Tapia.
Auxiliares: Obed Ulloa, Gustavo Navarro.

P1.

P2. **(a)**

(b) Una función $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ se dice concava-convexa si para cada $x \in \mathbb{R}^n$, la función $g_x : \mathbb{R}^n \to \mathbb{R}$ definida por $g_x(y) = f(x, y)$ es concava, y para cada $y \in \mathbb{R}^n$, la función $h_y : \mathbb{R}^n \to \mathbb{R}$ definida por $h_y(x) = f(x, y)$ es convexa.

I) (1.5 pts.) De una condición de segundo orden para que una función $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, dos veces diferenciable, sea concava-convexa.

Respuesta 1: Sabemos que para funciones 2 veces diferenciables, que el Hessiano sea semi definido positivo en todo punto es equivalente a la convexidad(0.8 pts), por esto, una condición de segundo orden para que f sea concava-convexa es:

- El Hessiano $H_{g_x}(y)$ es semi definido negativo, para todo $x,\ y \in \mathbb{R}^n$.
- El Hessiano $H_{h_y}(x)$ es semi definido positivo, para todo $x,\ y \in \mathbb{R}^n$.(0.7 pts ambas)

Respuesta 2: Notando que Hessiano de f es de la forma:

$$H_f(x, y) = \begin{pmatrix} A(x, y) & B(x, y) \\ B(x, y) & C(x, y) \end{pmatrix}, (0{,}5pts)$$

Donde $A(x, y), B(x, y), C(x, y)$ son matrices de n por n. Más aún, la matriz A contiene las segundas derivadas en x, C las segundas derivadas en y (y B las derivadas cruzadas). Por esto, f es concava-convexa si $A(x, y)$ es semi definida negativa para todo $x,\ y \in \mathbb{R}^n$ y $C(x, y)$ es semi definido positivo para todo $x,\ y \in \mathbb{R}^n$. (1.0 ptp)

II) (1.5 pts.) Suponga que $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ es una función concava-convexa y que $\nabla f(\overline{x}, \overline{y}) = 0$. Pruebe que:

$$f(\overline{x}, y) \leq f(\overline{x}, \overline{y}) \leq f(x, \overline{y}),\ \forall x,\ y \in \mathbb{R}^n.$$

Respuesta: Notando que el vector columna $\nabla f(x, y) = (\nabla h_y(x)^t, \nabla g_x(y)^t)^t$, entonces $\nabla h_{\overline{y}}(\overline{x}) = 0$ y $\nabla g_{\overline{x}}(\overline{y}) = 0$ (0.5 pts). Usando la concavidad (y convexidad) de $g_{\overline{x}}$ (y $h_{\overline{y}}$ respectivamente), obtenemos que $\overline{y}$ es un máximo globlal de $g_{\overline{x}}$ y que $\overline{x}$ es un mínimo global de $h_{\overline{y}}$ (0.5 pts). Con esto concluimos que:

$$f(\overline{x}, y) = g_{\overline{x}}(y) \leq g_{\overline{x}}(\overline{y}) = f(\overline{x}, \overline{y}) = h_{\overline{y}}(\overline{x}) \leq h_{\overline{y}}(x) = f(x, \overline{y}),\ \forall x,\ y \in \mathbb{R}^n.(0{,}5pts.)$$

P3. **(a)** Considere la función $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$, donde $U = \{(x, y) \in \mathbb{R}^2 :\ x > 0,\ y > 0\}$, definida por $f(x, y) = xy + 1/xy$.

I) (1 pto.) Estudie las condiciones necesarias y suficientes de optimalidad para la función f.
Respuesta: Veamos la condición de primer orden: $\nabla f(x, y) = 0$. Esto es:

$$\nabla f(x, y) = \begin{pmatrix} y - \dfrac{1}{yx^2} \\ x - \dfrac{1}{xy^2} \end{pmatrix} = 0.(0{,}2pts.)$$

Como $(x,y) \in U$, entonces $\nabla f(x,y) = 0$ si y sólo si $xy = 1$ (0.2 pts.). Para la condición de segundo orden calculamos el Hessiano de f.

$$H_f(x,y) = \begin{pmatrix} \dfrac{2}{yx^3} & 1 + \dfrac{1}{x^2y^2} \\ 1 + \dfrac{1}{x^2y^2} & \dfrac{2}{xy^3} \end{pmatrix} .(0{,}2pts.)$$

La primera entrada $2/yx^3$ es siempre positiva, luego para que el Hessiano $H_f(x,y)$ sea definido positivo, su determinante debe ser positivo.

$$\det(H_f(x,y)) = \frac{4}{x^4y^4} - (1 + \frac{1}{x^2y^2})(1 + \frac{1}{x^2y^2}),$$

que evaluada en los puntos críticos da 0, luego es sólo semi definida positiva (0.2 pts.). Por lo que no se puede concluir que son ni mínimos ni máximos locales (0.2 pts.).

II) (1 pto.) Muestre que el conjunto de puntos $H = \{(x,y) : \ xy = 1\}$ son mínimios globales de f.
Respuesta: Reagrupando la expresión y completando cuadrado, obtenemos:

$$f(x,y) = xy + \frac{1}{xy} = \frac{1 + x^2y^2}{xy} = \frac{(1 - xy)^2}{xy} + 2 \geq 2.$$

Como en los puntos críticos obtenemos que $f(x,y) = 2$, entonces corresponden a mínimos globales.(1.0 pto.)

III) (1 pto.) Considere ahora la función $g(x,y) = f(x,y) + (x-a)^2$, estudie la condición suficiente de optimalidad para g en función de $a \in \mathbb{R}$.
Respuesta: El gradiente de g es:

$$\nabla g(x,y) = \begin{pmatrix} y + \dfrac{-1}{yx^2} + 2(x-a) \\ x + \dfrac{-1}{xy^2} \end{pmatrix}, (0{,}2pts.)$$

el que si impones que sea 0, obtenemos que los puntos críticos satisfacen:

$$x^2y^2 - 1 + 2(x-a)xy^2 = 0, \qquad x^2y^2 = 1,$$

Por lo que $xy = 1$ y $x = a$ (0.2 pts.). Así, si $a > 0$ tenemos que g posee único punto crítico, en $(a, 1/a)$, mientras que si $a \leq 0$, g no posee puntos críticos. (0.2 pts.)

Veamos la matriz Hessiana de g en $(a, 1/a)$ cuando $a > 0$:

$$H_g(a,1/a) = \begin{pmatrix} \dfrac{2}{x^3y} + 2 & 1 + \dfrac{1}{x^2y^2} \\ 1 + \dfrac{1}{x^2y^2} & \dfrac{2}{xy^3} \end{pmatrix}_{(x,y)=(a,1/a)} = \begin{pmatrix} 2 + 2/a^2 & 2 \\ 2 & 2a^2 \end{pmatrix}, (0{,}2pts.)$$

que es definida positiva, por tanto $(a, 1/a)$ es un mínimo local.(0.2 pts.)

(b) **I)** (1.5 pts.) Sea la función $f : \mathbb{R}^2 \to \mathbb{R}$ definida por $f(x,y) = x^2 + 2y^2$ y sea $(x_0, y_0) \in \mathbb{R}^2$ distinto de $(0,0)$. Encuentre el mínimo de f y muestre que el algoritmo del gradiente partiendo en (x_0, y_0) converge en una sola iteración si y sólo si $x_0 = 0$ o $y_0 = 0$.
Respuesta: Como $f(x,y) \geq 0$, entonces el mínimo de f se encuentra en $(0,0)$ (0.2 pts.).

Sea $\vec{x}_0 = (x_0, y_0)$ y calculemos el siguiente vector usando el método del gradiente. Primero el gradiente:

$$\nabla f(\vec{x}_0) = (2x_0,\ 4y_0)^t.(0{,}2pts.)$$

definimos $g(\lambda) = f(\vec{x}_0 - \lambda \nabla f(\vec{x}_0))$ y buscamos su mínimo:

$$\begin{aligned} g(\lambda) &= (x_0 - 2\lambda x_0)^2 + 2(y_0 - 4\lambda y_0)^2 \\ g'(\lambda) &= 4x_0^2(2\lambda - 1) + 16y_0^2(4\lambda - 1).(0{,}2pts.) \end{aligned}$$

Con lo que concluímos que $g'(\overline{\lambda}) = 0$ se resuelve en:

$$\overline{\lambda} = \frac{x_0^2 + 4y_0^2}{2x_0^2 + 16y_0^2}.(0{,}2pts.)$$

Con esto, el vector $(x_1, y_1) := (x_0, y_0) - \overline{\lambda} \nabla f(x_0, y_0)$ está dado por:

$$(x_1, y_1) = \left(x_0 \left(1 - 2\frac{x_0^2 + 4y_0^2}{2x_0^2 + 16y_0^2} \right), y_0 \left(1 - 4\frac{x_0^2 + 4y_0^2}{2x_0^2 + 16y_0^2} \right) \right).(0{,}2pts.)$$

Si $x_0 = 0$, entonces es fácil ver que $x_1 = 0$ e $y_1 = 0$, luego hemos convergido en una iteración. Si $y_0 = 0$, el argumento es el mismo.(0.2 pts.)
Veamos la reciproca, supongamos que (x_1, y_1) es el vector 0, luego:

$$x_0 \left(1 - 2\frac{x_0^2 + 4y_0^2}{2x_0^2 + 16y_0^2} \right) = 0.$$

Si $x_0 = 0$, entonces ya llegamos a lo que queríamos probar, si $x_0 \neq 0$, entonces $8y_0^2/(2x_0^2 + 16y_0^2)$ es igual a 0, por lo que $y_0 = 0$.(0.3 pts.)

II) (1.5 pts.) Sea $(x_k, y_k)_k$ secuencia obtenida iterando el método del gradiente en f. Suponga que $x_0 > 0$ e $y_0 > 0$, pruebe que $x_k > 0$ para todo $k \in \mathbb{N}$, mientras que $y_k \neq 0$ para todo $k \in \mathbb{N}$ y va alternando de signo (esto es, y_{k+1} tiene el signo opuesto de y_k). **Indicación:** Utilice inducción.

Respuesta: Caso base: $x_0 > 0$ e $y_0 > 0$.(0.3 pts.)
Paso inductivo: Supongamos que $x_k > 0$ e $y_k \neq 0$. Usando los calculos realizados en la parte P3 $b)I$, obtenemos que:

$$(x_{k+1}, y_{k+1}) = \left(x_k \left(1 - 2\frac{x_k^2 + 4y_k^2}{2x_k^2 + 16y_k^2} \right), y_k \left(1 - 4\frac{x_k^2 + 4y_k^2}{2x_k^2 + 16y_k^2} \right) \right), (0{,}3pts.)$$

de donde podemos notar que:

$$1 - 2\frac{x_k^2 + 4y_k^2}{2x_k^2 + 16y_k^2} > 0, \qquad 1 - 4\frac{x_k^2 + 4y_k^2}{2x_k^2 + 16y_k^2} < 0,\ \forall (x_0, y_0) \neq (0, 0), (0{,}3pts.)$$

con lo que $x_k + 1$ sigue siendo mayor que 0 (0.3 pts.), e $y_k + 1$ es y_k múltiplicado por un número negativo, luego cambia de signo y sigue siendo distinto de 0.(0.3 pts.)

Duración: 3:00 hrs.