Universidad de Chile
Facultad de Ciencias Físicas y Matemáticas
Departamento de Ingeniería Matemática
31 de Mayo de 2011

Pauta Auxiliar Nº10

Profesor: Erwin Topp.
Auxiliares: Gianfranco Liberona, David Salas.

P1. El plano $x+y+2z=2$ y el paraboloide $z=x^2+y^2$ se intersectan en una elipse. Encuentre el punto más cercano y más lejano de este punto a origen.

Solución:

Consideremos el problema de optimización para encontrar el mínimo:

$$\begin{cases} & \min\limits_{x\in\mathbb{R}^3} \|x\| \\ s.a. & x+y+2z-2=0 \\ & z-(x^2+y^2)=0 \end{cases}$$

Como $(\cdot)^2$ es una función creciente en $\mathbb{R}_+$, el punto x^* donde f alcanza el mínimo, es el mismo punto donde f^2 alcanza el mínimo, si f es una función positiva. Luego, el problema anterior es equivalente a

$$\begin{cases} & \min\limits_{x\in\mathbb{R}^3} \|x\|^2 \\ s.a. & x+y+2z-2=0 \\ & z-(x^2+y^2)=0. \end{cases}$$

Escribamos entonces el Lagrangeano:

$$L(x,\lambda,\mu)=\|x\|^2+\lambda(x+y+2z-2)+\mu(z-x^2-y^2).$$

Entonces, tenemos que

$$\nabla_x L(x,\lambda,\mu)=\begin{pmatrix}2x\\2y\\2z\end{pmatrix}-\lambda\underbrace{\begin{pmatrix}1\\1\\2\end{pmatrix}}_{\nabla g}-\mu\underbrace{\begin{pmatrix}-2x\\-2y\\1\end{pmatrix}}_{\nabla h}.$$

Notemos que $\{\nabla g(x,y,z),\nabla h(x,y,z)\}$ es un conjunto linealmente independiente, salvo en los puntos $(-1/4,-1/4,z_0)$, con el $z_0\in\mathbb{R}$. Sin embargo, estos puntos no están en el conjunto factible (donde estamos buscando los óptimos). Es fácil chequearlo, basta con evaluar en las restricciones. Con esto, tenemos que en todo punto del conjunto $\{\vec{x}\in\mathbb{R}^3\ :\ g(\vec{x})=0,\ h(\vec{x})=0\}$, los vectores $\{\nabla g(x,y,z),\nabla h(x,y,z)\}$ son l.i. y podemos ocupar el teorema de Lagrange.

Tenemos entonces del sistema $\nabla_x L=0$, que

$$(1+\mu)x=(1+\mu)y$$

Si $\mu=-1$, entonces tendríamos que necesariamente $\lambda=0$ y luego $z=-1/2$. Pero esto no es posible por la restricción (2) $[z=x^2+y^2\geq 0]$. Luego, $\mu\neq -1$ y tendremos que $x=y$. Así, tenemos el sistema

$$\begin{cases} z=2x^2 \\ z=1-x \end{cases}$$

y luego $2x^2+x-1=0$, lo que implica que

$$x=\begin{cases} \frac{1+\sqrt{1+8}}{4} & =1 \\ \frac{1-\sqrt{1+8}}{4} & =-1/2 \end{cases}.$$

Evaluando, tenemos los puntos

$$P_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \qquad P_2 = \begin{pmatrix} -1/2 \\ -1/2 \\ 1/2 \end{pmatrix}$$

y $\|P_1\|^2 = 6 > 3/4 = \|P_2\|^2$, por lo que P_2 es el mínimo. Con el mismo procedimiento, P_1 es el máximo, pues solo hay 2 puntos críticos.

P2. Consideremos $\mathbb{R}^n$ con la norma euclidiana. Sea $A \in \mathcal{M}_{nn}(\mathbb{R})$ una matriz simétrica (i.e. $A = A^t$). Sea entonces la función

$$\begin{aligned} f : \mathbb{R}^n &\to \mathbb{R} \\ \vec{x} &\mapsto x^t A x \end{aligned}$$

(a) Demuestre que, $\forall x \in \mathbb{R}^n$, se tiene que

$$\nabla f(\vec{x}) = 2Ax$$

(b) Sea $S = \{x \in \mathbb{R}^n \ : \ \|x\| = 1\}$ (esfera unitaria en $\mathbb{R}^n$). Muestre que $x_0 \in \mathbb{R}^n$ es punto crítico de f restringida a S si y sólo si x_0 es vector propio de A normalizado.

(c) Suponga ahora que A es definida positiva (i.e. $\forall x \in \mathbb{R}^n \setminus \{0\},\ x^t A x > 0$). Sea $z \in \mathbb{R}^n$ no nulo, determine los puntos críticos de la función

$$\begin{aligned} g : \mathbb{R}^n &\to \mathbb{R} \\ \vec{x} &\mapsto \langle z, x \rangle \end{aligned}$$

restringida a $\mathcal{Q} = \{\vec{x} \in \mathbb{R}^n \ \ f(\vec{x}) = 1\}$.

Solución:

(a) Tenemos el siguiente desarrollo:

$$\begin{aligned} \frac{d}{dx_k}(x^t A x) &= \frac{d}{dx_i}\left(\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j\right) \\ &= \frac{d}{dx_i}\left(\sum_{j\neq k} a_{kj}x_k x_j + \sum_{i\neq k} a_{ik}x_i x_k + a_{kk}x_k^2 + \sum i,j \neq k a_{ij}x_i x_j + \right) \\ &= \sum_{j\neq k} a_{kj}x_j + \sum_{i\neq k} a_{ik}x_i + 2a_{kk}x_k \\ &= \sum_{j=1}^{n} a_{kj}x_j + \sum_{i=1}^{n} a_{ik}x_i \\ &= (Ax)_k + (A^t x)_k \\ &= (2Ax)_k \end{aligned}$$

donde $(Ax)_k$ es la coordenada $k-$ésima del vector Ax. Con esto, tenemos que

$$\nabla(x^t A x) = 2Ax.$$

(b) Notemos primero que $S = \{x \in \mathbb{R}^N \ : \ \|x\| = 1\} = \{x \in \mathbb{R}^N \ : \ \|x\|^2 = 1\}$. Entonces, el Lagrangeano que necesitamos considerar es

$$L(x, \lambda) = x^t A x - \lambda \|x\|^2$$

Luego,

$$\begin{aligned}
\nabla_x L(x,\lambda) &= 2Ax - \lambda\nabla(\|x\|^2) \\
&= 2Ax - \lambda\nabla(x^t x) \\
&= 2Ax - \lambda\nabla(x^t I_N x) \\
&= 2Ax - 2\lambda I_N x \\
&= 2(A - \lambda I_N)x
\end{aligned}$$

donde I_N es la matriz identidad. Luego,

$$\begin{aligned}
x_0 \text{es punto crítico} &\iff \nabla_x L(x_0,\lambda) = 0 \wedge \|x_0\|^2 = 1 \\
&\iff (A - \lambda I_N)x_0 = 0 \wedge \|x_0\| = 1 \\
&\iff x_0 \text{ es vector propio normalizado de } A.
\end{aligned}$$

(c) Como La restricción es solo una función, tenemos l.i. trivialmente. Luego

$$L(x,\lambda) = g(x) - \lambda(f(x) - 1)$$

y por lo tanto,

$$\nabla_x L(x,\lambda) = \nabla g(x) - \lambda\nabla f(x).$$

Es fácil notar que $\nabla g(x) = z$ y luego, los puntos críticos vienen dados por la ecuación

$$z - 2\lambda Ax = 0 \Longrightarrow x = \frac{1}{2\lambda}A^{-1}z$$

Reemplazando en la restricción, tenemos el siguiente desarrollo,

$$\begin{aligned}
f(x) = 1 &\iff \left(\frac{1}{2\lambda}A^{-1}z\right)^t A\left(\frac{1}{2\lambda}A^{-1}z\right) = 1 \\
&\iff \frac{1}{4\lambda^2}(A^{-1}z)^t z = 1 \\
&\iff \frac{1}{4}z^t(A)^{-1}z = \lambda^2 \qquad [\text{ por simetría de A }] \\
&\iff \lambda = \pm\left(\frac{1}{2}\sqrt{z^t(A)^{-1}z}\right)
\end{aligned}$$

Donde la raíz está bien definida, pues A es definida positiva y $z \neq 0$. Con esto, tenemos solo dos puntos críticos (dados por los valores de λ), y luego, uno es el máximo y el otro el mínimo.

P3. El objetivo de este problema es demostrar que la media geométrica es siempre menor que la media aritmética. Es decir,

$$\left(\prod_{i=1}^{N} x_i\right)^{1/N} \leq \frac{1}{N}\sum_{i=1}^{N} x_i.$$

Para esto, demuestre primero que basta considerar solo los casos donde $\sum_{i=1}^{N} x_i = 1$.

Solución:

Primero, notemos que

$$\underbrace{\left(\prod_{i=1}^{n}\frac{|x_i|}{\|x\|_1}\right)^{1/N} - \frac{1}{N}\sum_{i=1}^{N}\frac{|x_i|}{\|x\|_1}}_{(1)} = \underbrace{\frac{1}{\|x\|_1}\left(\left(\prod_{i=1}^{n}|x_i|\right)^{1/N} - \frac{1}{N}\sum_{i=1}^{N}|x_i|\right)}_{2}.$$

Por ende, si vemos que para todo elemento con $\|x\| = 1$ se cumple la desigualdad que queremos probar, entonces (1) sería menor o igual a 0 y luego (2) también lo sería. Finalmente eso implicaría que la desigualdad que queremos probar se tendría para todo punto $x \in \mathbb{R}^N$.

Podemos entonces considerar el problema como

$$(P) = \begin{cases} & \text{máx} \left(\prod_{i=1}^{N} |x_i|\right)^{1/N} - \frac{1}{N} \\ s.a. & \|x\|_1 = 1 \end{cases}$$

que es equivalente a

$$(P') = \begin{cases} & \text{máx} \left(\prod_{i=1}^{N} |x_i|\right)^{1/N} \\ s.a. & \|x\|_1 = 1 \end{cases}$$

Notemos que en este segundo problema, los mínimos se alcanzan cuando al menos una de las coordenadas es 0. Consideremos el cambio de variables, $\ln(\cdot)$. Esta función es creciente, por lo que los puntos donde se alcanzan los máximos y mínimos locales se conservan. El problema quedaría como

$$(P'') = \begin{cases} & \text{máx} \frac{1}{N} \left(\sum_{i=1}^{N} \ln(|x_i|)\right) \\ s.a. & \|x\|_1 = 1 \end{cases}$$

Notemos que en este caso, la función no es continua en todos los puntos (pues cuando algún x_i tiende a 0, la función tiene a $-\infty$. Sin embargo, sin contar estos puntos, los máximos locales deberían seguir cumpliendo el teorema de Lagrange. Luego,

$$\frac{d}{dx_i} L(x, \lambda) = (\frac{1}{N} \frac{1}{|x_i|} - \lambda)\text{sgn}(x_i).$$

Igualando $\nabla_x L(x, \lambda) = 0$, se tiene que

$$\forall i = 1, \dots, n, \ \frac{1}{|x_i|} = \lambda N,$$

o sea, todos los $|x_i|$ son iguales. El máximo se alcanza entonces en $|x_i| = \frac{1}{N}$ (esto se concluye de la restricción $\|x\|_1 = 1$). Luego,

$$\begin{aligned} \left(\prod_{i=1}^{N} |x_i|\right)^{1/N} &\leq \left(\prod_{i=1}^{N} \frac{1}{N}\right)^{1/N} \\ &= \frac{1}{N} = \frac{1}{N} \sum_{i=1}^{N} |x_i| \end{aligned}$$

con esto, tenemos que para todo $(x_1, \dots, x_n)$ con todos las coordenadas positivas, se tiene que

$$\left(\prod_{i=1}^{N} x_i\right)^{1/N} \leq \frac{1}{N} \sum_{i=1}^{N} x_i.$$