MA22A-**Cálculo en Varias Variables**
Profesor: Manuel del Pino
Auxiliares: Juan Campos y Julio Backhoff
Clase auxiliar 10

1. Se desea construir un estanque cilíndrico para almacenar bencina. Para la construcción del estanque debe considerarse la superficie del cilindro (costado) y las dos tapas.

   El material para la superficie del estanque se vende a 5 u.t.m. por metro cuadrado y el material para las tapas a 3 u.t.m. por metro cuadrado.

   Sabemos que el estanque se vacía cada mes y que para cada recarga mensual, donde debemos llenar completamente el estanque, solo contamos con 400 u.t.m. para comprar bencina. El precio de la bencina es 0.6 u.t.m./litro.

   (a) Determinar el conjunto de pares (r,h) (i..e. radio y altura) que cumplen con la máxima capacidad presupuestaria mensual.

   (b) Determinar el o los pares óptimos (r,h) que minimizan el costo de construcción del estanque, cumpliendo con la restricción de la parte (a.)

2. Sea $A \in \mathbb{R}^{nxn}$ una matriz simétrica real. Encuentre la solución de los problemas (1) y (2). Justifique la optimalidad de los puntos críticos.

$$(1) \quad \max_{||x||=1} x^t A x$$
$$(2) \quad \min_{||x||=1} x^t A x$$

3. La función de entropía $H(x) = -\sum_{i=1}^{n} x_i \log(x_i)$,definida para $x_i > 0$, tiene diversas aplicaciones en ingeniería, que van desde la codificación de mensajes hasta lo compresión de datos. En la *Teoría de la Información* la función de entropía representa una medida de incertidumbre asociada a una distribución de probabilidad $x = (x_i)_{i=1}^{n}$ de un mensaje emitido, es decir, la probabilidad de que un mensaje m sea igual a $i \in \{1, .., n\}$ es x_i (i.e. $Probabilidad(m = i) = x_i$). Así entre mayor sea el valor de $H(x)$ mayor será la incertidumbre asociada a la distribución x.

   (a) Estudie el máximo de la función de entropía en su dominio. Justifique que efectivamente se obtiene un máximo. Si ahora se redefine H para $x_i \geq 0$ considerando su comportamiento límite en 0 dado por $0log(0) = 0$, verifique que el máximo sigue siendo el mismo.

   (b) Muestre que la distribución de probabilidad que tiene máxima incertidumbre viene dada por $x_i = \frac{1}{n}$ $\forall i = 1, .., n$. ¿Es esto lógico?.

   NOTA: una distribución de probabilidad $x = (x_i)_{i=1}^{n}$ es una colección de números reales no negativos que satisfacen $\sum_{i=1}^{n} x_i = 1$.

Pauta

1. La superficie de un cilindro de radio r y altura h incluyendo las tapas es $S(r,h) = 2\pi r^2 + 2\pi rh$

   Llamemos mc =costo del material del costado= $5(\frac{utm}{m^2}) = 5 * 10^{-4}(\frac{utm}{cm^2})$ y mt =costo del material de las tapas= $3(\frac{utm}{m^2}) = 3 * 10^{-4}(\frac{utm}{cm^2})$

   (a) Como el estanque (cilindro) DEBE ser llenado completamente con las 400 utm que tenemos:
   V =volumen del cilindro= $400 * 0.6 = 240(litros) = 2.4 * 10^5(cc)$, pero el volumen de un cilindro es πhr^2, así que se tiene que los pares (r,h) que satisfacen la máxima capacidad presupuestaria mensual son los que cumplen:

   $$\pi r^2 h = 2.4 * 10^5 \quad (1)$$

   (b) Notemos que el costro de construir el estanque es $mc * 2\pi rh + mt * 2\pi r^2$. Además, la ecuación (1) la escribiremos como $hr^2 = K$, con $K = \frac{2.4*10^5}{\pi}$. Así, debemos resolver el siguiente problema de minimización con restricciones:

   $$\begin{array}{c} Minimizar\ mc * 2\pi rh + mt * 2\pi r^2 \\ sujeto\ a\ hr^2 = K \end{array}$$

   Así, debemos usar multiplicadores de Lagrange, notando que la restricción se escribe como $hr^2 - K = 0$:
   $L(r,h,\lambda) = mc * 2\pi rh + mt * 2\pi r^2 - \lambda(hr^2 - K) = arh + br^2 - \lambda(hr^2 - K)$ es el Lagrangiano, con $a = mc * 2\pi$ y $b = mt * 2\pi$
   Así, imponemos $\nabla L = 0$, quedando que :

   $$\begin{array}{c} \frac{\partial L}{\partial r} = ah + 2br - 2hr\lambda = 0 \\ \frac{\partial L}{\partial h} = ar - \lambda r^2 = 0 \\ hr^2 = K \end{array} \quad \Rightarrow \quad \begin{array}{c} ah + 2br = 2hr\lambda \ \ (2) \\ ar = \lambda r^2 \ \ (3) \end{array}$$

   Haciendo $\dfrac{(2)}{(3)}$ resulta $ah + 2br = 2ah$, de donde

   $$h = \frac{2br}{a} \quad (4)$$

   y recordando que $hr^2 = K$ y reemplazando (4) resulta:

   $$r = \sqrt[3]{\frac{aK}{2b}}$$

   Finalmente reemplazando los datos numéricos, $r = \sqrt[3]{\frac{2*10^5}{\pi}} = 39.929(cm)$ y $h = 39.929 * \frac{6}{5} = 47.915(cm)$

2. Como A es simétrica y real, sus valores propios son todos número reales. Como $||x|| = 1 \Leftrightarrow ||x||^2 = 1$, ocuparemos esta última restricción, la que es lo mismo que decir $x^t x = 1$.

Recordemos que $\nabla(x^tBx) = (B+B^t)x$ y por lo tanto si B es simétrica esto es igual a $2Bx$. Así el Lagrangeano del sistema es:

$$L(x,\lambda) = x^tAx - \lambda(x^tx - 1) = x^t(A-\lambda I)x + \lambda$$

donde I es la identidad de n por n. Así, como sabemos que un punto crítico del problema de optimización con restricciones debe satisfacer $\nabla L(x,\lambda) = 0$ se tiene:

$$\begin{aligned} 2(A-\lambda I)x &= 0 \\ ||x|| &= 1 \end{aligned}$$

Pero sabemos que la primera ecuación es resuelta solo por los vectores propios de A, en cuyo caso λ debe ser el valor propio correspondiente. Así, los puntos críticos del problema son los vectores propios (normalizados) de A.

Ahora, sea (x,λ) un vector propio de A (normalizado) y su valor propio. Así, $Ax = \lambda x$ y luego premultiplicando por x^t se obtiene que $x^tAx = \lambda x^tx = \lambda$. El valor de nuestra función objetivo sobre los puntos críticos es exactamente el valor propio del punto crítico. Luego, el mínimo y el máximo de x^tAx sujeto a $||x|| = 1$ es respectivamente el menor y el mayor valor propio de A, a saber, $\lambda_{\min}$ y $\lambda_{\max}$. Veamos que efectivamente, los vectores propios asociados a estos valores propios, que llamaremos $v_{\min}$ y $v_{\max}$ son óptimos.

Como A es simétrica, existe una base de $\mathbb{R}^n$ formada por vectores propios normalizados $(v_1, ..., v_n)$ todos ortogonales entre sí (es decir una base ortonormal) , y por lo tanto todo x en $\mathbb{R}^n$ se escribe como

$$x = a_1v_1 + ... + a_nv_n$$

Así $x^tx = (a_1v_1^t + ... + a_nv_n^t)(a_1v_1 + ... + a_nv_n)$. Pero $v_i^tv_j = 1$ si $i = j$ y 0 si no. Así, $x^tx = a_1^2 + ... + a_n^2 = ||x||^2$.

Ahora, $x^tAx = \sum_{i=1}^n \sum_{j=1}^n a_ia_jv_i^tAv_j = \sum_{i=1}^n \sum_{j=1}^n a_ia_jv_i^t\lambda_jv_j = \sum_{i=1}^n \sum_{j=1}^n a_ia_j\lambda_jv_i^tv_j = \sum_{i=1}^n a_i^2\lambda_i$

Así, $\lambda_{\min} \sum_{i=1}^n a_i^2 \leq x^tAx \leq \lambda_{\max} \sum_{i=1}^n a_i^2$ y así si $||x|| = 1$, queda que $\lambda_{\min} \leq x^tAx \leq \lambda_{\max}$, lo que demuestra la optimalidad de $v_{\min}$ y $v_{\max}$.

3. (a) La función está bien definida pues el dominio del logaritmo es exactamente los reales positivos.

   Notar que $\dfrac{\partial H}{\partial x_i} = -\log(x_i) - 1$, pues $x_i > 0$, y así $\dfrac{\partial^2 H}{\partial x_j \partial x_i} = 0$ si $i \neq j$ y $= -\dfrac{1}{x_i}$ si $i = j$. Así la matriz Hessiana de H es diagonal y todos sus términos en la diagonal (i.e., sus valores propios) son de la forma $-\dfrac{1}{x_i}$ y por lo tanto estrictamente negativos, luego la función H es cóncava (es decir tiene forma de "$\cap$") y por lo tanto tiene un máximo.

Por otra parte recordar que $\lim_{d\to 0} d\log(d) = 0$ y así podemos definir $H(0) = 0$ por continuidad. Notando que si todos los x_i son nulos salvo uno que sea por ejemplo igual a 0.5, H valdrá un número estrictamente positivo y por lo tanto 0 no puede ser máximo.

(b) Sea $x = (x_i)_{i=1}^n$ tal que $\sum_{i=1}^n x_i = 1$. Sea el Lagrangiano $L = H(x) - \lambda(\sum_{i=1}^n x_i - 1)$ e impongamos $\nabla L = 0$. Así resulta que:

$$-\log(x_i) - 1 - \lambda = 0 \quad ,\forall i$$

Así, $\log(x_i) = -1 - \lambda \quad \forall i$, y como "$-1 - \lambda$ es un número fijo" se tendrá que $\log(x_1) = \log(x_2) = ... = \log(x_n)$, lo que implica que todos los x_i son iguales, pero como $\sum_{i=1}^n x_i = 1$ la única forma de que esto ocurra es con $x_i = \dfrac{1}{n} \quad ,\forall i$.

En realidad esta respuesta es lógica pues si la probabilidad de que un mensaje dado sea igual a i es x_i, entonces la mayor incerteza se obtiene si todos los mensajes tienen la misma probabilidad.