IN3701 - Modelamiento y Optimización

Auxiliar Nº 12 - Optimización no lineal

Profesores: Daniel Espinoza, Fernando Ordoñez.
Auxiliares: Camilo Levenier, Carlos Bonet, Cristóbal Beltrán, Diego Fuentealba, Tomás Lagos, Sebastián Morales.

P1. Sea $Q \in \mathbb{R}^{n^2}$ una matriz cuadrada simétrica e invertible, con al menos un valor propio estrictamente negativo y $c \in \mathbb{R}^n$ un vector cualquiera. Consideremos el siguiente problema de optimización cuadrático:

$$(P)\ \min_{x\in\mathbb{R}^n} \quad f(x) = \frac{1}{2}x^tQx + c^tx$$

1. Encuentre una expresión resumida para $f(x+y)$, con $x, y \in \mathbb{R}^n$.

2. Encuentre un vector que sea dirección de decrecimiento asegurado para f, ie, un vector $v \in \mathbb{R}^n$ tal que:

$$\lim_{\theta\to\infty} f(x+\theta v) = -\infty$$

   ¿Qué se puede decir de (P)?

3. Suponga que Q fuese definida positiva estrictamente. ¿Sería posible encontrar una dirección de decrecimiento asegurado? ¿Qué se podría decir de (P)?

Hint: Recordar, una matriz cuadrada $A \in \mathbb{R}^{n^2}$ tiene como valor propio (o eigenvalue) a $\mu \in \mathbb{R}$ ssi $\exists v \in \mathbb{R}^n$ tal que $Av = \mu v$.

Pauta Pregunta 1:

1. Desarrollando la definicion de $f(x)$:

$$\begin{aligned} f(x+y) &= \frac{1}{2}(x+y)^tQ(x+y) + c^t(x+y) \\ &= \frac{1}{2}x^tQx + \frac{1}{2}x^tQy + \frac{1}{2}y^tQx + \frac{1}{2}y^tQy + c^tx + c^ty \\ &= f(x) + x^tQy + f(y). \end{aligned}$$

   Donde se utilizó que $x^tQy = y^tQx$ pues Q es simétrica.

2. Candidateamos a $v \in \mathbb{R}^n$, vector propio asociado a $\mu \in \mathbb{R}^n$, el valor propio negativo, de forma que $Qv = \mu v$. De esta forma tenemos:

$$\begin{aligned} f(x+\theta v) &= f(x) + \theta x^tQv + f(\theta v) \\ &= f(x) + \theta\mu x^tv + \frac{1}{2}\theta^2 v^tQv + \theta c^tv \\ &= f(x) + \theta(\mu x^tv + c^tv) + \frac{1}{2}\theta^2\mu v^tv \\ &= A\theta^2 + B\theta + C. \end{aligned}$$

   Donde $A = \frac{1}{2}\mu v^tv < 0$ pues $\mu < 0$ (es el valor propio negativo) y $v^tv > 0$.
   Entonces $f(x+\theta v)$ es un polinomio de segundo grado, el coeficiente asociado a su termino cuadratico nos da los límites. Como $A < 0$, se cumple que $\lim_{\theta\to\infty} f(x+\theta v) = \lim_{\theta\to\infty} A\theta^2 + B\theta + C = -\infty$. Luego v es dirección de dececimiento asegurado.

3. Sin entrar en cálculos, como Q es simétrica, sus vectores propios son linealmente independientes (de hecho son ortogonales), por lo que todo vector puede expresarse como combinación lineal de sus vectores propios. Como ninguno de estos vectores nos entrega una dirección de decrecimiento asegurado (por tener valores propios positivos), tampoco lo hará la combinación, por lo que el problema tendrá un mínimo global donde $\nabla f = 0$.

P2.

Para un vector $\alpha \in \mathbb{R}^n$, encuentre una solución para el siguiente problema mediante KKT.

$$\begin{array}{rll} \max \min\limits_{i=1,\dots,n} & (\alpha_i + x_i) & \\ s.t. & \sum_{i=1}^n x_i = 1 & \\ & x_i \geq 0 & i = 1, \dots, n \end{array}$$

Pauta Pregunta 2:
Consideremos la siguiente formulación alternativa:

$$\begin{array}{rll} \min & -z & \\ s.t. & z - \alpha_i - x_i \leq 0 & \\ & \sum_{i=1}^n x_i - 1 = 0 & \\ & -x_i \leq 0 & i = 1, \dots, n \end{array}$$

Este es problema lineal. Sabemos que dualidad fuerte se cumple. Las condiciones de KKT son necesarias y suficientes. El Lagrangeano es

$$z + \sum_{i=1}^n (\lambda_i(z - \alpha_i - x_i) + \mu x_i - \nu_i x_i) + \mu.$$

Las condiciones de KKT son:

$$\begin{aligned} \sum_{i=1}^n \lambda_i &= 1 \\ -\lambda_i + \mu - \nu_i &= 0 \\ \lambda_i(z - \alpha_i - x_i) &= 0 \\ \mu(1 - \sum_{i=1}^n x_i) &= 0 \\ \nu_i x_i &= 0 \\ \alpha_i + x_i &\geq z \\ \sum_{i=1}^n x_i &= 1 \\ x_i &\geq 0 \end{aligned}$$

Vemos que la solución debiese hacer multiples $x_i + \alpha_i = z$, mientras que el resto de los x_i debiesen ser 0. Confirmemos la intuición.

Caso 1: $\mu = 0$ En este caso vemos que $\lambda_i = \nu_i = 0$ para todo i. Sin embargo esto no es posible dado que $\sum_I \lambda_i = 1$. Concluimos que este caso no es posible.

Caso 2: $\mu > 0$ Supongamos que $x_i > 0$, entonces $\nu_i = 0$ y tenemos que $\lambda_i = \mu > 0$, por lo que $x_i + \alpha_i = z$. Por otro lado, si $x_i = 0$, entonces, $\alpha_i \geq z$.

Entonces, supongamos sin perdida de generalidad que $\alpha_1 \leq \alpha_2 \leq \ldots, \leq \alpha_n$ y definamos

$$i^* := \max\{i = 1, \ldots, n : \sum_{k=1}^{i} (\alpha_i - \alpha_k) \leq 1\}.$$

Se puede revisar que

$$\begin{aligned} z &= \frac{1 - \sum_{i=1}^{i^*} (\alpha_{i^*} - \alpha_i)}{i^*} \\ x_i &= \begin{cases} z - \alpha_i & i \leq i^* \\ 0 & i > i^* \end{cases} \\ \lambda_i &= \begin{cases} 1/i^* & i \leq i^* \\ 0 & i > i^* \end{cases} \\ \nu_i &= \begin{cases} 0 & i \leq i^* \\ 1/i^* & i > i^* \end{cases} \end{aligned}$$

satisface las condiciones de KKT.

P3.

Sea $f : \mathbb{R}^n :\mapsto \mathbb{R}$ una función convexa.

(i) Demuestre que para f, cualquier mínimo local es un mínimo global.

(ii) Pruebe que si f es estrictamente convexa, el mínimo, en caso de existir, es único.

(iii) Pruebe que si $x^* \in \mathbb{R}^n$ es tal que $\nabla f(x^*) = 0$, entonces x^* es un mínimo de f.

- **Obs:** Lo anterior quiere decir que para funciones convexas, la condición necesaria de primer orden es también suficiente.

Pauta Pregunta 3:

(i) Supongamos que x^* es óptimo local pero no global. Entonces se puede encontrar un punto $z \in \mathbb{R}^n$ con $f(z) < f(x^*)$. Consideremos el segmento convexo entre x^* y z esto es,

$$x = \lambda z + (1 - \lambda) x^*$$

para algún $\lambda \in (0, 1]$.
Como f es convexa, se tiene que

$$f(x) \leq \lambda f(z) + (1 - \lambda) f(x^*) < f(x^*)$$

Pero toda vecindad en torno a x^* contiene parte del segmento convexo descrito anteriormente, es decir, podemos tomar λ cercano a 1 de manera que $\|x - x^*\| \leq \epsilon$ para algún $\epsilon > 0$. Así, se tiene que x^* no es un mínimo local.

(ii) Sabemos que si f es convexa todo mínimo local es mínimo global, por lo que ambos son equivalentes. Supongamos que existen dos mínimos globales denotados por x_1 y x_2, cuyo valor denotamos como z. De la convexidad estricta de f se tiene que:

$$f\left(\frac{1}{2}x_1 + \frac{1}{2}x_2\right) < \frac{1}{2}f(x_1) + \frac{1}{2}f(x_2) = z$$

Lo que contradice la optimalidad de x_1 o x_2.

(iii) Como f es convexa se tiene que:

$$f(x) \geq f(x*) + \nabla f(x^*)^t (x - x^*) \quad \forall x, x^* \in \mathbb{R}^n$$

Como $\nabla f(x^*) = 0$ de lo anterior se tiene que

$$f(x) \geq f(x^*) \quad \forall x \in \mathbb{R}^n$$

Que corresponde a la definición de mínimo para f evaluado en x^*.