MA37A Optimización. Semestre Otoño 2008
Profesor: Héctor Ramírez C. **Auxiliares:** Omar Larré, Oscar Peredo.

Trabajo dirigido

1. Condiciones de optimalidad de primer orden

Consideremos una función $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ diferenciable. Recordemos que el cono de direcciones tangentes en x^* es $T_C(x^*) = \{d \in \mathbb{R}^n : \exists (\alpha_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}, (d_n)_{n\in\mathbb{N}} \subseteq \mathbb{R}^n; \alpha_n \to 0, d_n \to d, x^* + \alpha_n d_n \in C\}$.

1.1. Condiciones de optimalidad de primer orden caso general

Supongamos que $x^* \in C$ es un mínimo local de f. Entonces $\forall d \in T_C(x^*), \nabla f(x^*)d \geq 0$

1.2. Condiciones de optimalidad de primer orden: C convexo

Si C es convexo entonces $T_C(x^*) = \overline{\mathbb{R}_+(C - x^*)}$. Luego, si $x^* \in C$ es un mínimo local de f entonces $\forall x \in C, \nabla f(x^*)(x - x^*) \geq 0$.

1.3. Problema convexo

Si C es convexo y f es una función convexa, entonces $x^* \in C$ es un mínimo local de f si y sólo si $\forall x \in C, \nabla f(x^*)(x - x^*) \geq 0$.

Nota: En todos los casos, si x^* pertenece al interior de C, entonces $T_C(x^*) = \mathbb{R}^n$ y por lo tanto la condición de primer orden es que $\nabla f(x^*) = 0$.

2. Ejercicios

P1. Considere $f : \mathbb{R}^n \to \mathbb{R}$ diferenciable y tal que $\lim_{||x||\to\infty} f(x) = +\infty$. Muestre que f tiene un mínimo global. En particular, estudie la función $g(x) = \frac{1}{2}x^tQx + b \cdot x$, donde $b \in \mathbb{R}^n$ y Q es una matriz definida positiva.

SOL.:

Consideremos $C = f^{-1}(-\infty, f(0)+1]$. Es claro que C es no vacío. Además, como f es continua C es un cerrado, y como $\lim_{||x||\to\infty} f(x) = +\infty$ entonces C es acotado. Luego C es compacto, entonces f alcanza su mínimo x^* en C. Fuera de C todos las imágenes de f son mayores a cualquier punto en C, por lo tanto x^* es un mínimo global.

Para el caso en que $g(x) = \frac{1}{2}x^tQx + b \cdot x$, como Q es definida positiva, entonces $x^tQx \geq \lambda ||x||^2$ donde $\lambda > 0$ es el menor valor propio de Q (esta propiedad se puede demostrar con propiedades típicas de álgebra lineal, sin embargo la probaremos más adelante). Asi $\lim_{||x||\to\infty} g(x) = +\infty$

y por lo tanto g tiene un mínimo global $x^* \in \mathbb{R}^n$. Luego, la condición de primer orden implica que

$$\nabla g(x^*) = Qx^* + b = 0$$

y por lo tanto $x^* = -Q^{-1}b$.

P2. Sea $h : [0,1] \to [0,+\infty)$ una función continua. Estudie la existencia y caracterización de mínimos de

$$f(x_1,x_2) = b_1x_1 + b_2x_2 + c_1\int_0^{x_1} h(x)dx + c_2\int_{x_1}^{x_1+x_2} h(x)dx + c_3\int_{x_1+x_2}^{1} h(x)dx$$

sujeta a que $x_1, x_2 \geq 0$ y que $x_1 + x_2 \leq 1$. Las constante c_i, b_i son todas positivas.

SOL.:

Notemos que las restricciones definen un conjunto C convexo compacto (un triángulo en el plano, con vértices en (0,0), (1,0) y (0,1)). La función f es diferenciable y está restringida a un compacto, por lo tanto tiene un mínimo global $x^* = (x_1^*, x_2^*)$. Si asumimos que x^* está en el interior de C, luego $\nabla f(x^*) = 0$, es decir

$$\begin{aligned} b_1 + (c_1-c_2)h(x_1^*) + (c_2-c_3)h(x_1^*+x_2^*) &= 0 \\ b_2 + (c_2-c_3)h(x_1^*+x_2^*) &= 0 \end{aligned}$$

Veamos otros casos, donde el óptimo no está en el interior de C:

Caso $x_1^* = 0$, y $x_2^* \in (0,1)$

Supongamos que $x_1^* = 0$, y $x_2 \in (0,1)$. La condición de primer orden es $\forall x \in C, \nabla f(x^*)(x - x^*) \geq 0$, pero en este caso $(x - x^*)$ es un vector que es su primera componente sólo puede tener valores ≥ 0, mientras que en la segunda componente puede tener valores negativos y positivos. Por lo tanto en este caso la condición de primer orden es:

$$\begin{aligned} b_1 + (c_1-c_2)h(x_1^*) + (c_2-c_3)h(x_1^*+x_2^*) &\geq 0 \\ b_2 + (c_2-c_3)h(x_1^*+x_2^*) &= 0 \end{aligned}$$

Caso $x_1^* = 0$, y $x_2^* \in (0,1)$

Supongamos que $x_2^* = 0$, y $x_1 \in (0,1)$. Por un argumento análogo al anterior, la condición de primer orden es:

$$\begin{aligned} b_1 + (c_1-c_2)h(x_1^*) + (c_2-c_3)h(x_1^*+x_2^*) &= 0 \\ b_2 + (c_2-c_3)h(x_1^*+x_2^*) &\geq 0 \end{aligned}$$

Caso $x_1^* > 0$, y $x_2 + x_1^* = 1$

No es difícil ver que $(-1,1)$ y $(1,-1)$ son direcciones tangentes en x^*. Luego la condición $\forall d \in T_C(x^*), \nabla f(x^*)d \geq 0$ implica

$$-[b_1 + (c_1-c_2)h(x_1^*) + (c_2-c_3)h(x_1^*+x_2^*)] + [b_2 + (c_2-c_3)h(x_1^*+x_2^*)] \geq 0$$

$$[b_1 + (c_1-c_2)h(x_1^*) + (c_2-c_3)h(x_1^*+x_2^*)] - [b_2 + (c_2-c_3)h(x_1^*+x_2^*)] \geq 0$$

y entonces

$$b_1 + (c_1 - c_2)h(x_1^*) - b_2 = 0$$

Por otro lado $(0,-1)$ tambíen es una dirección tangente (pero no lo es $(0,1)$). Luego

$$0 \cdot [b_1 + (c_1 - c_2)h(x_1^*) + (c_2 - c_3)h(x_1^* + x_2^*)] - [b_2 + (c_2 - c_3)h(x_1^* + x_2^*)] \geq 0$$

entonces

$$b_2 + (c_2 - c_3)h(x_1^* + x_2^*) \leq 0$$

Se concluye que en este caso las condiciones necesarias son:

$$\begin{aligned} b_1 + (c_1 - c_2)h(x_1^*) - b_2 &= 0 \\ b_2 + (c_2 - c_3)h(x_1^* + x_2^*) &\leq 0 \end{aligned}$$

Caso x^* es un vértice del triángulo

Las condiciones necesarias en este caso no entregan mucha información extra.

3. Condiciones de optimalidad de primer orden: Teorema de Karush-Kuhn-Tucker (KKT)

Consideremos el problema (P):

$$\begin{array}{cc} \text{mín} & f(x) \\ sa & g_i(x) \leq 0,\ i \in I \\ & h_j(x) = 0,\ j \in J \\ & x \in \mathbb{R}^n \end{array}$$

donde todas las funciones son $\mathcal{C}^1$.

Sin dar detalles sobre la definición de un punto regular, diremos que si un punto x^* es regular, entonces satisface las hipótesis necesarias para usar el teorema de KKT. Es importante tener en cuenta que la regularidad de un punto está relacionado con las restricciones, no con la función objetivo.

3.1. Teorema de KKT

Para el problema (P), sea x^* un mínimo local regular. Luego existe un vector $\lambda \in \mathbb{R}^{|J|}$ y un vector $\mu \in \mathbb{R}^{|I|}$ con $\mu \geq 0$ tal que

$$\nabla f(x^*) + \sum_{j \in J} \lambda_j \nabla h_j(x^*) + \sum_{i \in I} \mu_i \nabla g_i(x^*) = 0$$

$$\mu_i g_i(x^*) = 0,\ \forall i \in I$$

Si además el problema (P) es convexo, es decir, h_j son funciones lineales afines, g_i son funciones convexas y f es convexo, entonces la condición anterior es también suficiente para que x^* sea un mínimo global (si f es estrictamente convexo, entonces el mínimo es único).

3.2. Condiciones suficientes de regularidad

Hay dos condiciones que veremos para que x^* sea regular. Consideremos el conjunto de restricciones

$$\begin{aligned}&g_i(x) \leq 0,\ i \in I\\&h_j(x) = 0,\ j \in J\\&x \in \mathbb{R}^n\end{aligned}$$

donde todas las funciones son $\mathcal{C}^1$. Sea x^* un punto factible. Denotemos por $I(x^*) := \{i \in I : g_i(x^*) = 0\}$ el conjunto de las restricciones activas para x^*.

3.2.1. Independencia Lineal de Gradientes Activos (ILGA)

Sea x^* un punto factible, tal que $\{\nabla h_j(x^*)\}_{j\in J}$ y $\{\nabla g_i(x^*)\}_{i\in I(x^*)}$ son vectores linealmente independientes. Entonces x^* es regular.

3.2.2. Condición de Slater

Supongamos que el conjunto de retricciones es convexo, es decir, h_j son funciones lineales afines y g_i son funciones convexas. Si se satisface la condición de Slater:

$$\exists \bar{x} \in \mathbb{R}^n : g_i(\bar{x}) < 0\ \forall i \in I, \quad h_j(\bar{x}) = 0\ \forall j \in J$$

entonces todo punto factible es regular.

4. Más problemas

P3. Sea Q una matriz de $n \times n$ simétrica. Muestre que $x^tQx \geq \bar{\lambda}||x||^2,\ \forall x \in \mathbb{R}^n$, donde $\bar{\lambda} = \min\{\lambda : \lambda \text{ } es\ valor\ propio\ de\ Q\}$ y que la igualdad se alcanza con el vector propio asociado al valor propio $\bar{\lambda}$.

SOL.:

Consideremos el problema no convexo (P3)

$$\begin{array}{ll}\text{mín} & x^tQx\\ sa & x^tx = 1\\ & x \in \mathbb{R}^n\end{array}$$

Sabemos que (P3) tiene solución, pues el conjunto de restricciones es compacto y la función objetivo continua.

Definamos $h(x) = 1 - x^tx$ (notar que esta restricción equivale a decir que $||x|| = 1$). Aqui sólo hay una restricción de igualdad: $h(x) = 0$. Notar que $\nabla h(x) = -2x$, y por lo tanto es linealmente independiente si $x \neq 0$. En este caso, todo punto factible es distinto de cero, y entonces todo punto factible es regular (por ILGA). Entonces la solución x tiene que satisfacer las ecuaciones de KKT:

$$2Qx - \lambda 2x = 0$$

para algún $\lambda \in \mathbb{R}$, es decir, $Qx = \lambda x$.

Esta última ecuación caracteriza a los candidatos a óptimo: tienen que ser vectores propios (cuyo multiplicador KKT asociado es su valor propio). Como Q tiene una base de n vectores propios de normalizados, entonces tenemos que cada uno de estos es un candidato a mínimo (es decir, tenemos n candidatos). Sea x_k un vector propio normalizado y λ_k su valor propio real asociado, luego

$$x_k^t Q x_k = \lambda_k ||x_k||^2 = \lambda_k$$

Como estos son todos los candidatos a óptimo, se concluye que la solución a (P3) es el vector propio normalizado asociado al menor valor propio $\bar{\lambda}$.

Para concluir, notemos que lo anterior implica que para $x \neq 0$, $\frac{x^t Q x}{||x||^2} \geq \bar{\lambda}$, y entonces $x^t Q x \geq \bar{\lambda}||x||^2, \ \forall x \in \mathbb{R}^n$.

P4. Considere el problema

$$\begin{array}{ll} \text{mín} & 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 \\ sa & \qquad x_1^2 + x_2^2 \leq 5 \qquad (1) \\ & \qquad 3x_1 + x_2 \leq 6 \qquad (2) \\ & \qquad x \in \mathbb{R}^n \end{array}$$

Escriba las condiciones KKT y encuentre el mínimo.

SOL.:

El problema tiene solución, pues el conjunto de restricciones es compacto, y la función objetivo continua.

Notemos que las restricciones son convexas, asi con $\bar{x} = 0$ se satisface las condiciones de Slater, y entonces todos los puntos factibles son regulares.

Además, la función objetivo es estrictamente convexa. En efecto

$$f(x_1, x_2) := 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^t \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 10x_1 - 10x_2$$

y la matriz $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ es definida positiva.

La condición de KKT es

$$4x_1 + 2x_2 - 10 + 2\mu_1 x_1 + 3\mu_2 = 0$$
$$2x_1 + 2x_2 - 10 + 2\mu_1 x_2 + \mu_2 = 0$$
$$\mu_1 \geq 0, \ \mu_2 \geq 0$$
$$\mu_1(x_1^2 + x_2^2 - 5) = 0$$
$$\mu_2(3x_1 + x_2 - 6) = 0$$

Para encontrar una solución definimos varias combinaciones de restricciones activas. En este problema podemos probar con ninguna restriccion activa, una o dos restricciones activas.

Restricción (1) inactiva y (2) restricción activa

Suponiendo eso tenemos

$$4x_1 + 2x_2 - 10 + 2\mu_1 x_1 + 3\mu_2 = 0$$
$$2x_1 + 2x_2 - 10 + 2\mu_1 x_2 + \mu_2 = 0$$

$$3x_1 + x_2 = 6$$

y $\mu_1 = 0$ (pues (1) esta inactivo).

Despejando x_2

$$4x_1 + 2(6 - x_1) - 10 + 3\mu_2 = 0$$

$$2x_1 + 2(6 - x_1) - 10 + \mu_2 = 0$$

multiplicando por -2 la primera y sumando tenemos que $5\mu_2 = -2$, lo cual contradice que $\mu_2 \geq 0$. No tenemos solución al sistema KKT.

Restricción (1) activa y restricción (2) inactiva

Suponiendo eso tenemos

$$4x_1 + 2x_2 - 10 + 2\mu_1 x_1 = 0$$

$$2x_1 + 2x_2 - 10 + 2\mu_1 x_2 = 0$$

$$x_1^2 + x_2^2 = 5$$

y $\mu_2 = 0$ (pues (2) esta inactivo). El sistema anterior tiene la solución $x_1 = 1,\ x_2 = 2, \mu_1 = 1$. Luego, como el problema es convexo y como f es estrictamente convexa $x_1 = 1, x_2 = 2$ es el óptimo global único. Como es único, no seguimos buscando.