Optimización

Primer Cuatrimestre 2015

Práctica N° 2: Métodos de descenso.

Ejercicio 1 Dada una constante $b \in \mathbb{R}$, considerar la función punto a conjunto dada por:

$$f(x) = \{y \in \mathbb{R}^n : y^t x \leq b\}, \quad \forall x \in \mathbb{R}^n.$$

¿f es cerrada?

Ejercicio 2 Sean $f : X \to Y$ y $g : Y \to Z$ dos funciones punto a conjunto. Probar que si f es cerada en x, g es cerrada en $f(x)$ e Y es compacto, entonces $f \circ g$ es cerada en x.

Ejercicio 3 Sean $f : X \to Y$ punto a punto y $g : Y \to Z$ punto a conjunto. Probar que si f es continua en x y g es cerrada en $g(x)$, entonces $f \circ g$ es cerrada en x.

Ejercicio 4 Mostrar que si A es una aplicación punto a punto, en el Teorema de Convergencia Global puede eliminarse la hipótesis de que los puntos x_k caigan sobre un compacto.

Ejercicio 5 Sea $f : \mathbb{R}^n \to \mathbb{R}$. El algoritmo de descenso genérico consiste en, dado $x_k \in \mathbb{R}^n$, dar una dirección d_k y un paso t_k tal que $f(x_k + t_k d_k) < f(x_k)$, y tomar $x_{k+1} = x_k + t_k d_k$.

a) Probar que si f es C^1 y $\nabla f(x_k) \cdot d_k < 0$, entonces d_k es una dirección de descenso.

b) Concluir que $-\nabla f(x_k)$ es una dirección de descenso.

Ejercicio 6 Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función cuadrática, y sea d_k una dirección de descenso en el punto x_k. Probar que el paso óptimo está dado por:

$$t_k = -\frac{d_k^t \nabla f(x_k)}{d_k^t H f(x_k) d_k}.$$

Ejercicio 7 Una función $f : \mathbb{R} \to \mathbb{R}$ se dice *unimodal* en el intervalo $[a, b]$ si existe $x^* \in (a, b)$ tal que f es estrictamente decreciente en (a, x^*) y estrictamente creciente en (x^*, b). Probar que si f es unimodal, entonces, dados α, β tales que $a < \alpha < \beta < b$ vale que:

- Si $f(\alpha) \leq f(\beta)$, entonces f es unimodal en $[a, \beta]$,
- Si $f(\alpha) \leq f(\beta)$, entonces f es unimodal en $[\alpha, b]$.

Ejercicio 8 Dada una función f unimodal en $[a, b]$, se propone el siguiente algoritmo para buscar su mínimo x^*:

1. Se fijan $a_0 = a$, $b_0 = b$.
2. Para $k = 1, 2, \dots$ se eligen α_k, β_k tales que $a_k < \alpha_k < \beta_k < b_k$.
   2.1. Si $f(\alpha_k) \leq f(\beta_k)$, se toman: $a_{k+1} = a_k$, $b_{k+1} = \beta_k$.
   2.2. Si $f(\alpha_k) > f(\beta_k)$, se toman: $a_{k+1} = \alpha_k$, $b_{k+1} = b_k$

Probar que $\lim a_k = \lim b_k = x^*$. ¿Cuántas veces debe evaluarse f para calcular $[a_{n+1}, b_{n+1}]$?

Ejercicio 9 (Búsqueda por la razón dorada:) Se desea fijar un criterio para la elección de α_k, β_k en el algoritmo anterior, de manera tal que se cumplan:

- que en cada paso el intervalo se vea reducido en un factor fijo η:

$$\beta_{k+1} - \alpha_{k+1} = \eta(\beta_k - \alpha_k),$$

- que en cada paso sea necesario evaluar f una sola vez. Es decir, que alguno de los nuevos puntos: α_{k+1} ó β_{k+1} coincida con alguno de los anteriores α_k ó β_k.

Escribir las fórmulas para α_{k+1} y β_{k+1} en función de a_k, b_k y η para que se satisfaga la primer condición, y calcular el valor de η para que se cumpla la segunda.

Ejercicio 10 Implementar un algoritmo que reciba como entrada una función f, un intervalo $[a, b]$, y una tolerancia δ y calcule el mínimo de f en $[a, b]$ con error menor o igual que δ, mediante el algortimo de búsqueda por la razón dorada.

Ejercicio 11 Dada $\delta > 0$, sea la función punto a conjunto $\mathbf{S}^\delta$ definida como

$$\mathbf{S}^\delta(x, d) = \Big\{ y : y = x + \alpha d, \quad 0 \leq \alpha \leq \delta; \quad f(y) = \min_{0 \leq \beta \leq \delta} f(x + \beta d) \Big\}.$$

Explicar lo que hace $\mathbf{S}^\delta$ y probar que si f es continua entonces $\mathbf{S}^\delta(x, d)$ es cerrada en (x,d). ¿Por qué es importante este resultado?

Ejercicio 12 Sea $\varepsilon > 0$, sea la función punto a conjunto $\mathbf{S}^\varepsilon$ definida como

$$\mathbf{S}^\varepsilon(x, d) = \Big\{ y : y = x + \alpha d, \quad 0 \leq \alpha; \quad f(y) \leq \min_{0 \leq \beta} f(x + \beta d) + \varepsilon \Big\}.$$

Explicar lo que hace $\mathbf{S}^\varepsilon$ y probar que si f es continua y $d \neq 0$ entonces $\mathbf{S}^\varepsilon(x, d)$ es cerrada en (x,d). ¿Por qué es importante este resultado?

Ejercicio 13 (Búsqueda Compacta)

Sea $f : \mathbb{R}^n \to \mathbb{R}$ una función C^2, *estríctamente* convexa con un único mínimo x^* y tal que $f(x) \to +\infty$ cuando $||x|| \to +\infty$.

a) Probar que, fijado un $x_0 \in \mathbb{R}^n$, el conjunto

$$\{y \in \mathbb{R}^n : f(y) < f(x_0)\}$$

es acotado.

b) Sea $\mathcal{D} = \{e_1, -e_1, e_2, -e_2, \ldots, e_n, -e_n\}$. Probar que para todo $x \neq x^*$ existe $d \in \mathcal{D}$ dirección de descenso en x. (Es decir, existe $d \in \mathcal{D}$ y $t_0 > 0$ tal que $f(x + td) < f(x)\ \forall\, t \in (0, t_0)$).

c) Consideramos, para $x \in \mathbb{R}^n$, $\alpha > 0$, la función punto a conjunto

$$A(x, \alpha) = \{y \in \mathbb{R}^n : y = x + \alpha d,\, d \in \mathcal{D}\}$$

Probar que A es cerrada en (x, α) si $\alpha \neq 0$.

d) Estudiamos el algoritmo de Búsqueda Compacta, dado por:

i. Tomar $x_0 \in \mathbb{R}^n$, $\alpha_0 > 0$, $k = 0$.
ii. Si existe $y \in A(x_k, \alpha_k)$ tal que $f(y) < f(x)$: Poner $x_{k+1} = y$, $\alpha_{k+1} = \alpha_k$, $k = k + 1$.
Si no: Poner $\alpha_k = \frac{\alpha_k}{2}$
iii. Ir a *ii*.

Probar que $x_k \to x^*$ $(k \to \infty)$.

Ejercicio 14 Implementar un algoritmo que reciba como datos una función f, un punto x_k y una dirección d_k y aplique la condición de Armijo con Backtracking para determinar el paso del descenso, devolviendo el correspondiente x_{k+1}.

Ejercicio 15 Implementar un programa similar al del ejercicio anterior, pero utilizando la condición de Goldstein.

Ejercicio 16 Probar que la condición de Goldstein determina un algoritmo de búsqueda cerrado.

Ejercicio 17 (Braquistocrona) Considerar el problema de la *braquistocrona* consistente en hallar la curva que minimiza el tiempo de caída de una particula por efecto de la gravedad. Concretamente, buscamos una función $\varphi : [0, 1] \to [0, 1]$ tal que $\varphi(0) = 1$, $\varphi(1) = 0$ tal que minimice el funcional:

$$T(\varphi) = \int_0^1 \sqrt{\frac{1 + \varphi'(x)^2}{2g(1 - \varphi(x))}},$$

donde g es la aceleración gravitatoria. Para resolver este problema se realiza una discretización en la variable x: $0 = x_0 < x_1 < \ldots < x_n < x_{n+1} = 1$ y una discretización en φ: $1 = \varphi_0 > \varphi_1 > \ldots > \varphi_n > \varphi_{n+1} = 0$, de manera tal que φ_i será la aproximación de $\varphi(x_i)$. Por comodidad, asumiremos que la discretización en φ está fijada de antemano (por ejemplo: $\varphi_{i+1} - \varphi_i = \frac{1}{n+1}$), y que las x_i son las variables cuyo valor debe optimizarse.

a) Probar que si se asume que φ es lineal en los intervalos de la discretización, el funcional T puede escribirse:

$$\widetilde{T} = \sqrt{\frac{2}{g}} \sum_{i=0}^{n} \sqrt{1 + \Big(\frac{x_{i+1} - x_i}{\varphi_{i+1} - \varphi_i}\Big)^2} \Big(\sqrt{1 - \varphi_{i+1}} - \sqrt{1 - \varphi_i}\Big).$$

b Calcular analíticamente el gradiente de la expresión anterior de $\widetilde{T}$.

c Implementar funciones que reciban como input el número n de incógnitas de la discretización y devuelvan el funcional $\widetilde{T}$ y su gradiente.

d Calcular la curva braquistocrona minimizando el funcional $\widetilde{T}$ a través de los distintos métodos estudiados hasta el momento: método de Newton, método de búsqueda compacta y método del gradiente con búsqueda lineal por la razón de oro, y por búsqueda lineal inexacta utilizando la condición de Armijo y la condición de Goldstein. Comparar los resultados.

Ejercicio 18 Sea $Q \in \mathbb{R}^{nxn}$ definida positiva y sean $v_1, ...v_n$ vectores l.i. Mostrar que el método de Gram-Schmidt puede ser usado para generar una secuencia de direcciónes Q-ortogonales desde los v_i. Específicamente, muestre que

$$d_1 = v_1; \qquad d_{k+1} = v_{k+1} - \sum_{i=1}^{k} \frac{v_{k+1}^t Q d_i}{d_i^t Q d_i} d_i$$

forma un conjunto $Q-$ortogonal.

Ejercicio 19 Sea $f(x) = \frac{1}{2}x^t Q x - b^t x$ con Q DP. Sea x_1 un minimizante de f en un subespacio S_1 que contiene al vector d y sea x_2 un minimizante de f en un subespacio S_2 que contiene a d. Mostrar que si $f(x_1) < f(x_2)$ entonces $\overline{x} = x_1 - x_2$ es $Q-$ortogonal a d.

Ejercicio 20 Implementar el método del Gradiente Conjugado para minimizar una función cuadrática $f(x) = \frac{1}{2}x^t Q x - b^t x$:

(1) A partir de un x_0 tomar $d_0 = -g_0 = b - Qx_0$

(2) Para $k = 0, 1, ...n - 1$ hacer:

  (a) Hacer $x_{k+1} = x_k + \alpha_k d_k$ con

$$\alpha_k = \frac{-g_k^T d_k}{d_k^T H(x_k) d_k}, \qquad g_k = Qx_k - b.$$

  (b) Hacer $d_{k+1} = -g_{k+1} + \beta_k d_k$ con

$$\beta_k = \frac{g_{k+1}^T Q d_k}{d_k^T Q d_k}.$$

Ejercicio 21 Si definimos $\mathcal{B}_k =< d_0, \cdots, d_{k-1} >$ el subespacio generado por las primeras k direcciones conjugadas, mostrar que el método de las direcciones conjugadas, en cada x_k minimiza la función objetivo tanto en la recta $L : x_{k-1} + \alpha d_{k-1} : \alpha \in \mathbb{R}$, como en la variedad lineal $x_0 + \mathcal{B}_k$.

Ejercicio 22 Implementar el siguiente algoritmo que generaliza el del gradiente conjugado a funciones no cuadráticas:

(1) A partir de un x_0 tomar $g_0 = \nabla f(x_0)^T$ y hacer $d_0 = -g_0$.

(2) Para $k = 0, 1, ...n - 1$ hacer:

   (a) Hacer $x_{k+1} = x_k + \alpha_k d_k$ con $\alpha_k = \frac{-g_k^T d_k}{d_k^T H(x_k) d_k}$.

   (b) Hacer $g_{k+1} = \nabla f(x_{k+1})^T$.

   (c) Si $k \neq n - 1$, hacer $d_{k+1} = -g_{k+1} + \beta_k d_k$ con

$$\beta_k = \frac{g_{k+1}^T H(x_k) d_k}{d_k^T H(x_k) d_k}.$$

   y repetir (a).

(3) Hacer $x_0 = x_n$ y volver a (1).