Optimización

Primer Cuatrimestre 2018

Práctica N° 2: Métodos de descenso.

Algoritmos y convergencia

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 cerrada en el punto x, g es cerrada en el conjunto $f(x)$ e Y es compacto, entonces $g \circ f$ es cerrada 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 el punto x y g es cerrada en el punto $f(x)$, entonces $g \circ f$ es cerrada en x.

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

Ejercicio 5 (Orden de Convergencia) Sea $\{e_k\}$ una sucesión de números no negativos convergente a 0. Decimos que $\{e_k\}$ *converge linealmente* (o geométricamente) si existen $q > 0$ y $\beta \in (0, 1)$ tales que

$$e_k \leq q\beta^k, \qquad \forall k \geq 0.$$

Decimos que $\{e_k\}$ *converge superlinealmente* si para todo $\beta \in (0, 1)$ existe $q > 0$ tal que $e_k \leq q\beta^k$. Finalmente, dado $p > 1$, decimos que $\{e_k\}$ *converge al menos superlinealmente con orden p* si existen $q > 0$ y $\beta \in (0, 1)$ tales que $e_k \leq q\beta^{p^k}$ para todo $k \geq 0$.

a. Verificar que si para algún $\beta \in (0, 1)$ vale que

$$\limsup_{k\to\infty} \frac{e_{k+1}}{e_k} = \beta,$$

entonces $\{e_k\}$ converge linealmente.

b. Mostrar que si

$$\limsup_{k\to\infty} \frac{e_{k+1}}{e_k} = 0,$$

entonces $\{e_k\}$ converge superlinealmente.

c. Verificar que si $\{e_k\}$ converge superlinealmente con orden $p > 1$, entonces, efectivamente, converge superlinealmente.

d. Mostrar que si

$$\limsup_{k\to\infty} \frac{e_{k+1}}{e_k^p} = \beta,$$

entonces $\{e_k\}$ converge superlinealmente con orden p.

Métodos de descenso

Ejercicio 6 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 $t_k > 0$ 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 diferenciable en x_k 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 7 (Descenso más rápido) Sea f de clase C^1 en $\mathbb{R}^n$.

(a) Mostrar que la dirección de $-\nabla f(x_k)$ es la de máximo descenso. Es decir, mostrar que la pendiente de $\phi(t) := f(x_k + td)$ en $t = 0$ se minimiza entre todas las direcciones d de norma 1 en $d^* = -\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}$.

(b) **(Efecto zigzag)** Dada la sucesión $(x_k)_{k\geq 1}$ generada por el método del descenso más rápido con búsqueda lineal óptima, mostrar que las direcciones entre iteraciones consecutivas son ortogonales.

Ejercicio 8 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 9 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 y continua en $[a, b]$ entonces tiene un único mínimo en $[a, b]$. Probar además que 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 $f(\alpha) \geq f(\beta)$, entonces f es unimodal en $[\alpha, b]$.

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 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 para determinar el paso del descenso, devolviendo el correspondiente x_{k+1}.

Ejercicio 14 Implementar un programa similar al del ejercicio anterior, pero utilizando la condición de Goldstein.

Ejercicio 15 Probar que la condición de Goldstein determina un algoritmo de búsqueda cerrado.

Ejercicio 16 Sea $Q \in \mathbb{R}^{n\times n}$ 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 17 Sea $f(x) = \frac{1}{2}x^tQx - b^tx$ con Q simétrica 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 $\overline{x} = x_1 - x_2$ es $Q-$ortogonal a d.

Ejercicio 18 Implementar el método del Gradiente Conjugado para minimizar una función cuadrática $f(x) = \frac{1}{2}x^tQx - b^tx$:

(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

$$\alpha_k = \frac{-g_k^t d_k}{d_k^t Q d_k}, \qquad g_k = Qx_k - b.$$

   (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 19 Dada $f(x) = \frac{1}{2}x^tQx - b^tx$ con Q simétrica definida positiva, 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 20 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 $\alpha_k = \frac{-g_k^T d_k}{d_k^T Hf(x_k) d_k}$.

(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

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

y repetir (a).

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