Universidad de Buenos Aires - Facultad de Ciencias Exactas y Naturales - Depto. de Matemática

Optimización

Primer Cuatrimestre 2019

Práctica N° 2: Métodos de descenso

Ejercicio 1 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. Es decir, que existe $\delta > 0$ tal que $f(x_k + td_k) < f(x_k)$ para todo $t \in (0, \delta)$.

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

c) Probar que si f es diferenciable en x_k y d_k es una dirección de descenso en x_k, entonces $\nabla f(x_k)^t d_k \leq 0$.

Ejercicio 2 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 3 Considerar la función cuadrática $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \frac{1}{2}x^t Ax + b^t x + c$, siendo A simétrica y definida positiva. Demostrar que si un método de direcciones de descenso con búsqueda lineal exacta es utilizado para minimizar f entonces la longitud óptima del paso, a partir de un punto x en la dirección d, es $\lambda = -\frac{d^t \nabla f(x)}{d^t Ad}$.

Ejercicio 4 Una función $f : \mathbb{R} \to \mathbb{R}$ se dice *estrictamente 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 estrictamente unimodal y continua en $[a, b]$ entonces tiene un único mínimo en $[a, b]$.

Ejercicio 5 Considerar $h : \mathbb{R}^n \to \mathbb{R}$ una función cuadrática dada por $h(x) = \frac{1}{2}x^t Ax + b^t x + c$ con A simétrica y definida positiva. Sea x^* el minimizador de h y sea $f(x) = h(x + x^*) - h(x^*) = \frac{1}{2}x^t Ax$. Sea $(x_k)_k$ la sucesión generada por el método de gradiente con búsqueda exacta aplicado a f y sea $(y_k)_k$ dada por $y_k = x_k + x^*$. Demostrar que el método del gradiente con búsqueda exacta aplicado a h genera la sucesión $(y_k)_k$.

Ejercicio 6 Sea $f : \mathbb{R}^2 \to \mathbb{R}$ dada por $f(x, y) = \frac{1}{2}(x^2 - y)^2 + \frac{1}{2}(1 - x)^2$. ¿Cuál es el minimizador de f? Realizar una iteración del método de Newton para minimizar f a partir de $x_0 = (2, 2)$. ¿Es un buen paso? Calcular $f(x_0)$ y $f(x_1)$.

Ejercicio 7 Sea $Q \in \mathbb{R}^{n\times n}$ definida positiva y sean $v_1, \dots, v_n$ vectores l.i. Mostrar que el método de Gram-Schmidt puede ser usado para generar una secuencia de direcciones Q-ortogonales desde los v_i. Específicamente, mostrar 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 8 Considerar una función cuadrática $f : \mathbb{R}^n \to \mathbb{R}$ dada por $f(x) = \frac{1}{2}x^t Ax + b^t x + c$ con A simétrica y definida positiva. Sea $\mathcal{B}_k = \langle d_0, \dots, d_{k-1}\rangle$ 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 9 El criterio de Armijo exige un $t \in \mathbb{R}$ para el cual

$$\varphi(t) = f(x + td) \leq f(x) + \sigma\ t\ d^T \nabla f(x) = \varphi(0) + \sigma\ t\ \varphi^{'}(0)$$

con $\sigma \in (0, 1)$. Si f es una función cuadrática entonces φ es una parábola. Demostrar que si el minimizador t^* de la parábola cumple la condición, entonces $\sigma \in (0, \frac{1}{2})$.

Ejercicio 10 Sean $f : \mathbb{R}^n \to \mathbb{R}$, x, $d \in \mathbb{R}^n$, $\lambda > 0$ tales que $x + \lambda d$ cumple la condición de Armijo. Sea $0 < \mu < \lambda$. ¿Cumple μ la condición de Armijo? Demostrar o dar un contraejemplo.

Ejercicio 11 Considerar la función $f : \mathbb{R}^2 \to R$ dada por

$$f(x, y) = x - y + 2x^2 + 2xy + y^2.$$

a) Demostrar que $d = (-1, 0)$ es una dirección de descenso para f a partir de $(0, 0)$. Analizar cuál es el paso máximo que se puede dar en esa dirección para hacer decrecer el valor de f utilizando búsqueda exacta y utilizando Armijo.

b) Para la dirección de máximo decrecimiento en $(0, 0)$ determinar el intervalo de paso máximo que se puede dar en esa dirección para hacer decrecer el valor de f utilizando la regla de Armijo con parámtero $\sigma = \frac{1}{4}$.

Ejercicio 12 Sean $f : \mathbb{R}^n \to \mathbb{R}$, de clase C^2, $x \in \mathbb{R}^n$ tal que $\nabla f(x) = 0$ y $Hf(x)$ no es semidefinida positiva. Demostrar que existe una dirección de descenso a partir de x.

Ejercicio 13 En el proceso de minimizar una función $f : \mathbb{R}^n \to \mathbb{R}$, de clase C^1, el iterado x_k se obtuvo haciendo búsqueda lineal a lo largo de la dirección $d_{k-1} \neq 0$. Determinar una dirección d_k ortogonal a d_{k-1}, de descenso a partir de x_k que sea combinación lineal de d_{k-1} y $-\nabla f(x_k)$.

Ejercicio 14 Sea $f : \mathbb{R}^n \to \mathbb{R}$, de clase C^1. Para $k \in \mathbb{N}$ se define $x_{k+1} = x_k - \lambda_k \nabla f(x_k)$ donde $\lambda_k \geq \bar{\lambda} > 0$ para todo $k \geq 0$. Suponiendo que $\{x_k\}$ converge a x^*, demostrar que $\nabla f(x^*) = 0$.

Ejercicio 15 *(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, se busca 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: $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = 1$ y una discretización en φ: $1 = \varphi_0 > \varphi_1 > \cdots > \varphi_n > \varphi_{n+1} = 0$, 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 y método del gradiente con búsqueda lineal por la sección áurea, y por búsqueda lineal inexacta utilizando las condicies de Armijo y de Wolfe. Comparar los resultados.