Optimización (MAT-275) 2^{do} **Semestre 2009**

TAREA 1

Pedro Gajardo Rafael Plaza Muñoz

Instrucciones generales

- La tarea debe ser entregada impresa y enviada a los correos electrónicos plazarafael13@gmail.com y pedro.gajardo@usm.cl. La versión impresa puede ser dejada en la oficina F332 hasta el día 20 de Noviembre a las 18h00.
- La versión electrónica debe contener un archivo comprimido (e.g. tar o zip) con 7 carpetas (una por cada ejercicio). Cada una de ellas tendrá el nombre *ejerciciox* donde x se refiere al número del ejercicio en cuestión. En las carpetas deberán colocar los archivos ejecutables (MATLAB, SCILAB, u otros) utilizados para resolver los ejercicios.
- La versión impresa, debe ser escrita en LaTeX.
- La tarea es **individual**. Si comenta alguno de los problemas con sus compañeros, debe especificar y dar los créditos por las ideas entregadas en los resultados que entregue.

Problema 1

Sea $\{a_1, a_2, \ldots, a_n\} \subset \mathbb{R}^n$ un conjunto de vectores linealmente independientes y Q una matriz de $n \times n$, simétrica y definida positiva.

(a) Mostrar que los vectores $d_1, d_2, \ldots, d_n$ definidos por

$$d_k = \begin{cases} a_k & \text{si } k = 1 \\ a_k - \sum\limits_{i=1}^{k-1} \left(\frac{\langle d_i, Qa_k \rangle}{\langle d_i, Qd_i \rangle} \right) d_i & \text{si } k \geq 2, \end{cases}$$

son Q-conjugados.

(b) Obtener los vectores d_k a partir de $a_1 = (\ 1,\ \ 0,\ \ 0\ )$, $a_2 = (\ 1,\ \ -1,\ \ 4\ )$, $a_3 = (\ 2,\ \ -1,\ \ 6\ )$ y

$$Q = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 3 & 2 \\ -1 & 2 & 2 \end{pmatrix}.$$

(c) Definir un algoritmo para minimizar una función $f : \mathbb{R}^n \mapsto \mathbb{R}$ que demore a lo más n iteraciones si f es cuadrática, que haga uso del anterior método para obtener direcciones conjugadas. Para ello entregue un pseudo-código que haga uso del paquete algorithm2e de LaTeX (o sea deben haber **while**, **for**, **if**, etc.).

(d) Programe el código y aplíquelo al problema:

$$\min_{x \in \mathbb{R}^3} 2x_1^2 + 3x_1x_2 + 4x_2^2 + 2x_3^2 - 2x_2x_3 + 5x_1 + 3x_2 - 4x_3$$

utilizando búsqueda lineal exacta y de Wolfe. Considere como punto de partida el origen y al menos otros 9 puntos elegidos al azar. Compare los resultados con ambas búsquedas lineales y comente.

Problema 2

Considere el siguiente método de optimización sin restricciones:

- **Inicialización**: Escoger un nivel de tolerancia ϵ y un punto inicial x_1. Sea $y_1 = x_1$, $d_1 = -\nabla f(y_1)$ y $k = j = 1$.
- **Pasos**:

Paso 1) Sea λ_j el escalar obtenido por búsqueda lineal para minimizar la función $\lambda \mapsto f(x_j + \lambda d_j)$ e $y_{j+1} = y_j + \lambda_j d_j$. Si $j = n$ ir al Paso 4, de lo contrario ir al Paso 2.

Paso 2) Sea $d = -\nabla f(y_{j+1})$ y $\widehat{\mu}$ obtenido por búsqueda lineal para minimizar la función $\mu \mapsto f(y_{j+1}+\mu d)$. Sea $z_1 = y_{j+1} + \widehat{\mu} d$, $i = 1$ e ir al Paso 3.

Paso 3) Si $\|\nabla f(z_i)\| < \epsilon$, detener el algoritmo y entregar z_i. De lo contrario, sea μ_i el escalar obtenido al aplicar búsqueda lineal para minimizar la función $\mu \mapsto f(z_i + \mu d_i)$ y $z_{i+1} = z_i + \mu_i d_i$. Si $i < j$, reemplazar i por i+1 y repetir el Paso 3. Sino, definir $d_{j+1} = z_{j+1} - y_{j+1}$, luego reemplazar j por $j + 1$ e ir al Paso 1.

Paso 4) Sea $y_1 = x_{k+1} = y_{n+1}$, $d_1 = -\nabla f(y_1)$, j=1 y reemplazar k por k+1. Ir al Paso 1.

(a) Entregar un pseudo-código de este algoritmo usando el paquete *algorithm2e* de de LaTeX. Considere la búsqueda lineal como un subprocedimiento.

(b) Demostrar que este algoritmo con búsqueda lineal exacta, se detiene en a lo sumo n pasos cuando $f(x) = \langle x, Qx\rangle - \langle b, x\rangle$, con Q una matriz de $n \times n$ simétrica definida positiva.

(c) Programar el algoritmo y minimizar $(x_1 - 2)^4 + (x_1 - 2x_2)^2$, usando búsqueda lineal exacta tipo Newton e inexacta tipo Wolfe, comenzando del punto $(0, 3)$ ¿Se alcanza el mínimo global?. Compare los procedimientos y comente.

Problema 3

Sean Q y M_1 matrices de $n \times n$, simétricas y definidas positiva. Considere la función cuadrática

$$f(x) = \frac{1}{2}\langle x, Qx\rangle - \langle b, x\rangle.$$

El conjunto $\{x_1, x_2, \ldots, x_n\} \subset \mathbb{R}^n$ se define recursivamente por

- $x_1 \in \mathbb{R}^n$.
- λ_j con $j = 1, \ldots, n$ escalares obtenidos mediante búsqueda lineal exacta asociada a la función $\lambda \mapsto f(x_j + \lambda d_j)$, donde $x_{j+1} = x_j + \lambda_j d_j$ con $d_j = -M_j \nabla f(x_j)$ y M_{j+1} dada por:

$$\begin{aligned}
M_{j+1} &= M_j + \frac{(p_j - M_j q_j)(p_j - M_j q_j)^t}{\langle q_j, p_j - M_j q_j\rangle} \\
p_j &= x_{j+1} - x_j \\
q_j &= \nabla f(x_{j+1}) - \nabla f(x_j).
\end{aligned}$$

Asuma que el conjunto $\{p_1, \ldots, p_n\}$ es linealmente independiente.

(a) Demostrar que la matriz sumada a M_j para obtener M_{j+1} es de rango uno.

(b) Muestre que para $j = 1, \ldots, n$ se cumple $p_i = M_{j+1} q_i$ para $i \leq j$.

(c) Demostrar que $M_{n+1} = Q^{-1}$.

(d) Pruebe que las direcciones d_j son de descenso en x_j.

(e) Demostrar que M_{j+1} es definida positiva cuando M_j lo es, si y sólo si $\langle q_j, p_j - M_j q_j \rangle > 0$.

(f) Programe el algoritmo descrito y utilícelo para minimizar $x_1 - 4x_2 + 2x_1^2 + 2x_1x_2 + 3x_2^2$, utilizando búsqueda lineal exacta. Pruebe con diez vectores iniciales aleatorios.

Problema 4

Dada la función de mérito $q(t) = f(x + td)$ en el punto x con dirección de descenso d, considere la búsqueda lineal inexacta cuyo criterio de parada es

$$\sigma \leq \frac{q(t) - q(0)}{tq'(0)} \leq 1 - \sigma,$$

donde σ es un parámetro en $(0, 1/2)$. Asuma que en los intervalo de búsqueda $[t_L, t_R]$ el escalar $t \in [t_L, t_R]$ es elegido de tal forma que si se produce una infinidad de extrapolaciones, entonces $t_L \to +\infty$ y si se realiza una infinidad de interpolaciones entonces $|t_R - t_L| \to 0$.

(a) Establezca condiciones sobre la función f para que la búsqueda lineal descrita anteriormente se detenga en un número finito de iteraciones.

(b) Imponiendo alguna condición sobre la elección de las direcciones de descenso (d_k), como las vistas en clases, establezca resultados para obtener que la sucesión generada por $x_{k+1} = x_k + t_k d_k$, donde $t_k > 0$ es elegido mediante la búsqueda lineal descrita, satisface

$$\liminf_{k \to +\infty} \|\nabla f(x_k)\| \to 0.$$

Problema 5

Para el método del gradiente conjugado visto en clases, considere las siguientes formulas para β_{k-1}:

$$\beta_k = \frac{\langle g_k, g_k \rangle}{\langle g_{k-1}, g_{k-1} \rangle} \tag{1}$$

$$\beta_k = \frac{\langle g_k, g_k - g_{k-1} \rangle}{\langle d_{k-1}, g_k - g_{k-1} \rangle} \tag{2}$$

$$\beta_k = \frac{\langle g_k, g_k - g_{k-1} \rangle}{\langle g_{k-1}, g_{k-1} \rangle} \tag{3}$$

$$\beta_k = -\frac{\langle g_k, g_k \rangle}{\langle d_{k-1}, g_{k-1} \rangle} \tag{4}$$

$$\beta_k = \frac{\langle g_k, g_k \rangle}{\langle d_{k-1}, g_k - g_{k-1} \rangle} \tag{5}$$

donde d_k es la dirección de descenso obtenida por el método en el iterado x_k y $g_k = \nabla f(x_k)$ (f función objetivo a minimizar).

(a) Demuestre que si aplica el método del gradiente conjugado para minimizar la función $f(x) = \langle x, Qx \rangle - \langle b, x \rangle$ con Q una matriz de $n \times n$, simétrica, definida positiva entonces (1), (2), (3), (4) y (5) son iguales. De un ejemplo que permita demostrar que esto no ocurre cuando la función a minimizar no es cuadrática.

(b) Aplique el método del gradiente conjugado con búsqueda lineal de Wolfe fuerte y con la búsqueda lineal descrita en el Problema 4, para cada uno de los posibles valores de β_k sobre las siguientes funciones.

$$\begin{aligned} f_1(x) &= (1{,}5 - x_1 + x_1x_2)^2 + (2{,}25 - x_1 + x_1x_2^2)^2 + (2{,}625 - x_1 + x_1x_2^3)^2 \text{ con } -4{,}5 \le x_i \le 4{,}5, i = 1, 2. \\ f_2(x) &= x_1^2 + 2x_2^2 - 0{,}3\cos(3\pi x_1) - 0{,}4\cos(4\pi x_2) + 0{,}7 \text{ con } -100 \le x_i \le 100, i = 1, 2. \\ f_3(x) &= \left(x_2 - \frac{5}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{8}{\pi}\right)\cos(x_1) + 10 \text{ con } -5 \le x_1 \le 10, 0 \le x_2 \le 15. \end{aligned}$$

Para $f_1(x)$ considere como punto de partida todo punto en una *lattice* cuadrada de lado 0,5 sobre el dominio de la función. Para $f_2(x)$ considere cuadrados de lado 10 y para $f_3(x)$ de lado 5. Entregue una gráfica de cada una de las funciones y además de las trayectorias que produce el algoritmo para cada punto de partida. Compare y comente.

Problema 6

Sea $f : \mathbb{R}^n \mapsto \mathbb{R}$ una función estrictamente convexa. Considere los iterados $x_{k+1} = x_k + \alpha_k d_k$, donde d_k es una dirección de descenso en x_k. Asuma que existe $L > 0$ tal que,

$$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\| \text{ para todo } x, y \in \mathbb{R}$$

y existen escalares positivos c_1, c_2 tales que para todo k se cumple:

$$\begin{aligned} c_1\|\nabla f(x_k)\|^2 &\le -\langle \nabla f(x_k), d_k\rangle \\ \|d_k\|^2 &\le c_2\|\nabla f(x_k)\|^2. \end{aligned}$$

Si el paso α_k está determinado por:

$$\alpha_{k+1} = \begin{cases} \alpha_k & \text{si } \langle \nabla f(x_{k+1}), d_k\rangle \le 0 \\ \beta\alpha_k & \text{en otros casos,} \end{cases}$$

donde $\beta \in (0, 1)$ es un escalar fijo y $\alpha_0 > 0$ es cualquier escalar positivo. Entonces,

(a) Mostrar que el paso se reduce después de una iteración k-ésima si y sólo si el intervalo I_k conectando x_k y x_{k+1} contiene en su interior un vector $\overline{x}_k$ que minimiza $f(x)$ sobre $x \in I_k$.

(b) Pruebe que el paso será constante después de un número finito de iteraciones.

(c) Defina un algoritmo que haga uso del punto anterior y una dirección d_k que cumpla con las condiciones de este ejercicio. Aplíquelo para minimizar la función:

$$f(x) = \begin{cases} x^2\left(\sqrt{2} - \sin\left(\frac{5\pi}{6} - \sqrt{3}\ln(x^2)\right)\right) & \text{si } x \ne 0 \\ 0 & \text{si } x = 0, \end{cases}$$

considerando diferentes condiciones iniciales.

Problema 7

Sea $f : \mathbb{R}^n \mapsto \mathbb{R}$ una función continuamente diferenciable. Si x^* es un mínimo local de f a través de toda línea que pasa por x^*, es decir, la función

$$g(\alpha) = f(x^* + \alpha d)$$

es minimizada en $\alpha = 0$, para todo $d \in \mathbb{R}^n$. Entonces,

(a) Mostrar que $\nabla f(x^*) = 0$.

(b) Considerar la función $f(y, z) = (z - y^2)(z - 4y^2)$. Mostrar que el origen es un mínimo local de f a través de toda línea que pasa por $(0, 0)$, pero no es un mínimo local.