UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Departamento de Matemática
2007

Certamen 2

Curso: Optimización (MAT 275)
Profesor: Pedro Gajardo
Fecha: 5 de noviembre 2007

Pregunta 1

Considere el siguiente problema de minimización cuadrática

$$\min_{x\in\mathbb{R}^n} f(x) := \frac{1}{2}\langle x, Ax\rangle - \langle b, x\rangle, \tag{PC}$$

donde A es una matriz simétrica de $n \times n$ definida positiva y $b \in \mathbb{R}^n$. Dado un $x_0 \in \mathbb{R}^n$ se genera la sucesión

$$x_{k+1} = x_k + t_k d_k$$

donde d_k es la dirección de máximo descenso en x_k y t_k es obtenido a través de búsqueda lineal exacta.

1. Obtenga una expresión para los iterados x_{k+1} sólo en función de x_k.
2. Pruebe que
$$\frac{1}{2}\|x - x^*\|_A^2 = f(x) - f(x^*) \qquad \forall\, x \in \mathbb{R}^n$$
donde x^* es la solución de (PC) y $\|\cdot\|_A$ es la norma definida por la matriz A.
3. Muestre las siguientes igualdades:
$$\begin{aligned}
\|x_k - x^*\|_A^2 \quad - \quad \|x_{k+1} - x^*\|_A^2 &= 2t_k\langle \nabla f(x_k), A(x_k - x^*)\rangle - t_k^2\langle \nabla f(x_k), A\nabla f(x_k)\rangle \\
\|x_k - x^*\|_A^2 \quad - \quad \|x_{k+1} - x^*\|_A^2 &= \frac{\langle \nabla f(x_k), \nabla f(x_k)\rangle^2}{\langle \nabla f(x_k), A\nabla f(x_k)\rangle} \\
\|x_k - x^*\|_A^2 &= \langle \nabla f(x_k), A^{-1}\nabla f(x_k)\rangle. \\
\|x_{k+1} - x^*\|_A^2 &= \left\{1 - \frac{\langle \nabla f(x_k), \nabla f(x_k)\rangle^2}{\langle \nabla f(x_k), A\nabla f(x_k)\rangle\langle \nabla f(x_k), A^{-1}\nabla f(x_k)\rangle}\right\}\|x_k - x^*\|_A^2.
\end{aligned}$$
4. Demuestre que si el iterado inicial x_0 es tal que $x_0 - x^*$ es paralelo a un vector propio de la matriz A, entonces el método converge en una iteración.

Pregunta 2

1. Considere el problema

$$\min_{x\in\mathbb{R}^n} f(x) \tag{P}$$

donde $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función de clase $\mathcal{C}^1$ con gradiente Lipschitz. Suponga que el conjunto solución de (P), que escribiremos $S(P)$, es no vacío y notemos f^* al valor óptimo de (P).

   a) Muestre que existe una constante $c_1 > 0$ tal que

$$f(x) \leq f^* + c_1 \mathrm{d}^2(x, S(P)),$$

   donde $\mathrm{d}(\cdot, S(P))$ es la función distancia al conjunto $S(P)$.

   b) Si además la función f es convexa y existe una vecindad V del conjunto $S(P)$ tal que

$$f^* + \frac{l}{2}\mathrm{d}^2(x, S(P)) \leq f(x) \qquad \forall\, x \in V,$$

   para una cierta constante $l > 0$, pruebe que

$$l(f(x) - f^*) \leq 2\|\nabla f(x)\|^2 \qquad \forall\, x \in V.$$

   c) Demuestre que, bajo las hipótesis del punto anterior, dada una sucesión $\{x_k\}_{k\in I\!N}$ convergente a una solución de (P), las velocidades de convergencia de $f(x_k)$ a f^*, de x_k a $S(P)$ y de $\nabla f(x_k)$ a 0, son la misma. Para ello pruebe que existen constantes positivas p_1, p_2, p_3 y p_4 tales que, para todo x en una vecindad de $S(P)$, se tiene

$$\|\nabla f(x)\|^2 \leq p_1\|x - x^*\|^2 \leq p_2(f(x) - f^*) \leq p_3\|\nabla f(x)\|^2 \leq p_4\|x - x^*\|^2,$$

   donde $x^* \in S(P)$ es la proyección (¿por qué existe y es única?) de x sobre $S(P)$.

2. Una función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ se dice fuertemente convexa si existe $r > 0$ tal que $f - \frac{r}{2}\|\cdot\|^2$ es convexa.

   Sea $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ una función fuertemente convexa de clase $\mathcal{C}^1$ con gradiente Lipschitz y $\{x_k\}_{k\in I\!N}$ una sucesión generada por

$$x_{k+1} = x_k + t_k d_k,$$

   donde d_k es una dirección de descenso en x_k y t_k es obtenido a través de la búsqueda lineal de Wolfe. Demuestre que existen dos constantes positivas q_1 y q_2 tales que

$$q_1\|\nabla f(x_k)\| \cos\theta_k \leq \|x_{k+1} - x_k\| \leq q_2\|\nabla f(x_k)\| \cos\theta_k,$$

   donde θ_k es el ángulo entre la dirección de descenso d_k y $-\nabla f(x_k)$.

Pregunta 3

Considere el problema

$$\min_{x\in\mathbb{R}^n} f(x) \tag{P}$$

donde $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función convexa, acotada inferiormente y de clase $\mathcal{C}^1$.

Para cada $\lambda > 0$ se definen las funciones

$$f_\lambda(x) = \min_{y\in\mathbb{R}^n} f(y) + \frac{1}{2\lambda}\|y - x\|^2$$

y $P_f^\lambda : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ dada por $P_f^\lambda(x) = x_\lambda$ donde

$$f_\lambda(x) = f(x_\lambda) + \frac{1}{2\lambda}\|x_\lambda - x\|^2.$$

1. Justifique que las funciones f_λ y P_f^λ están bien definidas.
2. Muestre que $f_\lambda \leq f$ para todo $\lambda > 0$ y f_λ es decreciente con respecto a λ.
3. Pruebe que para todo $\lambda > 0$ se tiene

$$\langle P_f^\lambda(x) - P_f^\lambda(y), x - y\rangle \geq \|P_f^\lambda(x) - P_f^\lambda(y)\|^2 \qquad \forall\, x, y \in \mathbb{R}^n.$$

4. Dada una sucesión de números reales positivos $\{\lambda_k\}_{k\in\mathbb{N}}$ y un elemento inicial $x_0 \in \mathbb{R}^n$, se define la sucesión

$$x_{k+1} = P_f^{\lambda_k}(x_k).$$

   Demuestre que

   a) $f(x_k)$ es decreciente;

   b) Si x^* es una solución del problema (P), entonces la sucesión $\|x_k - x^*\|$ es decreciente;

   c) Si la sucesión λ_k es acotada inferiormente entonces $\nabla f(x_k)$ converge a cero.

   d) Si el conjunto solución del problema (P) es no vacío y la sucesión λ_k es acotada superiormente, entonces la sucesión x_k tiene al menos una subsucesión convergente y todo punto de acumulación pertenece al conjunto solución.

Tiempo: 3 horas.