Aspectos Matemáticos y Computacionales del Método de Mejora Máxima por Bloque

Juan Pablo Soto Quirós
jusoto@tec.ac.cr

Instituto Tecnológico de Costa Rica

12 de agosto del 2019

1. Problema de Optimización para una Función Multivariable

2. Mejora Máxima por Bloques

1 Problema de Optimización para una Función Multivariable

2 Mejora Máxima por Bloques

Función multivariable

Una función escalar de n variables $f : \mathbb{R}^n \to \mathbb{R}$, asigna a cada punto $(x_1, ..., x_n) \in \mathbb{R}^n$, un único número real denotado con $f(x_1, ..., x_n) \in \mathbb{R}$.

Problema

Sea $\mathbf{x} = (x_1, ..., x_n)$. El objetivo de esta presentación es explicar un conjunto de métodos iterativos que permitirán dar una solución al problema de minimización

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}).$$

- En resumen, los pasos para resolver el problema

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

en un curso de cáluclo en varias variables son los siguientes:

1. Calcular los puntos críticos (Gradiente).
2. Calcular el Hessiano Orlado (Hessiano Orlado).
3. Evaluar puntos críticos en el Hessiano Orlado.
4. Calcular determinantes de las submatrices principales.
5. Interpretar resultado.

- La solución del problema $\min\limits_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x})$ se realizará a través de métodos iterativos.

- Es decir, dado un valor inicial $\mathbf{x}^{(0)} \in \mathbb{R}^n$, cada método iterativo generará una sucesión de puntos

$$\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, ..., \mathbf{x}^{(k)}, ...\},$$

  donde $\mathbf{x}^{(j)} \in \mathbb{R}^n$, para todo $j = 1, 2, ...$

- La sucesión de puntos puede tener tres criterios de convergencia:

  1. La sucesión converge a la solución del problema.
  2. La sucesión converge a a un punto que no es solución del problema.
  3. La sucesión no converge.

1. Problema de Optimización para una Función Multivariable

2. Mejora Máxima por Bloques

- Un método iterativo de optimización por bloques es el método conocido como **mejora máxima por bloques** (MMB).
- El método MMB actualiza cada variable usando un enfoque tipo *greddy*.
- El método MMB actualiza el bloque de variables correspondiente al bloque de mejora máxima usando la regla de Jacobi, pero compara el error de cada una de las actualizaciones.
- Al final, la actualización con el error mínimo es el nuevo elemento de la sucesión

Algorithm 2 Método de Mejora Máxima por Bloques

Requires: $\mathbf{x}^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)}) \in \mathbb{R}^n$
Returns: $\mathbf{x}^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)}) \in \mathbb{R}^n$

for $k = 1, 2, ...$ **do**
    for $j = 1, ..., n$ **do**
        $\widetilde{x}_j^{(k)}$ = Usando la regla de Jacobi
        $e_j^{(k)} = f(x_1^{(k-1)}, \ldots, x_{j-1}^{(k-1)}, \widetilde{x}_j^{(k)}, x_{j+1}^{(k-1)}, \ldots, x_n^{(k-1)})$
    end
    $i = \arg \min_{j=1,\ldots,p} \left\{ e_j^{(k)} \right\}$
    for $j = 1, ..., n$ **do**
        if $j = i$ **then**
            $x_j^{(k)} = \widetilde{x}_j^{(k)}$
        else
            $x_j^{(k)} = x_j^{(k-1)}$
        end
    end
    if el criterio de parada se cumple **then**
        return $\mathbf{x}^{(k)} = (x_1^{(k)}, \ldots, x_n^{(k)})$
    end
end

Ejemplo 2

Considere la función

$$f(x, y, z) = (x-2)^2 + (y+3)^2 + (x+y+z)^2.$$

Aplicando el método MMB con un valor inicial $\mathbf{x}^{(0)} = (1, 1, 1)$, obtenemos los siguientes resultados.

k	$\mathbf{x}^{(k)}$	$f(\mathbf{x}^{(k)})$
0	$(1, 1, 1)$	30
1	$(1, -2.5, 1)$	5.5
2	$(1, -2.5, 0.25)$	4.375
3	$(1.75, -2.5, 0.25)$	2.125
4	$1.75, -2.5, (-0.125)$	1.8438
$\vdots$	$\vdots$	$\vdots$
27	$(2.5, -2.5, -0.5)$	1

- Una desventaja del método MMB es el exceso de cálculo, lo cual hace que el método converja más lento, comparado con el método OA.
- Sin embargo, esta desventaja se puede reducir si se implementa en paralelo.
- Otra ventaja es que este método es menos restrictivo, en términos de convergencia.

Convergencia del Método MMB

Si el conjunto $\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leqslant f(\mathbf{x}^{(0)})\}$ es un conjunto acotado, entonces la sucesión generada por el método MMB converge a un mínimo coordenado de f.