Hoja de Trabajo 6 - Gradient Descent Methods

Eduardo Andrés Santizo Olivet (21007361)

1. Dado el problema de optimización no restringida:

$$\min_{(x_1,x_2)\in\mathbb{R}^2} \quad f(x_1,x_2) = 4x_1^2 + 2x_2^2 + 4x_1x_2 - 3x_1.$$

   a) Aplique las SOSC para resolver dicho problema. *Ayuda:* notar que f es una función cuadrática.

   b) Aplique el algoritmo *gradient-descent* con *line-search exacto* para resolver dicho problema. Utilice como punto inicial $x_0 = (2,2)^T$ y realice *"a mano"* únicamente 3 iteraciones.

A. $f(x_1, x_2) = 4x_1^2 + 2x_2^2 + 4x_1x_2 - 3x_1$

$$\nabla f(x_1,x_2) = \begin{bmatrix} 8x_1 + 4x_2 - 3 \\ 4x_2 + 4x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- $\left.\begin{array}{l} 8x_1 + 4x_2 - 3 = 0 \\ 4x_2 + 4x_1 = 0 \end{array}\right\}$ $x_2 = -x_1$
- $8x_1 - 4x_1 - 3 = 0$
- $4x_1 = 3$
- $x_1 = \frac{3}{4} \quad x_2 = -\frac{3}{4}$

PUNTO ESTACIONARIO: $\begin{bmatrix} \frac{3}{4} & -\frac{3}{4} \end{bmatrix}$

$$\nabla^2 f(x_1,x_2) = \begin{bmatrix} 8 & 4 \\ 4 & 4 \end{bmatrix}$$

$\lambda_1 = 2(3+\sqrt{5}) = 10.47$

$\lambda_2 = 2(3-\sqrt{5}) = 1.53$

Obtenido con Wolfram

La matriz Hessiana es positiva definida, por lo que se puede establecer que se cumplen tanto las FONC y las SOSC. Además, dado que se encontró un punto estacionario (Y el problema es de carácter convexo), se puede establecer que el mismo consiste del minimizador global del problema.

B. $k = 0: \quad x_0 = [2, 2]^T$

$$\nabla f(x_0) = \begin{bmatrix} 8x_1 + 4x_2 - 3 \\ 4x_2 + 4x_1 \end{bmatrix} = \begin{bmatrix} 8(2) + 4(2) - 3 \\ 4(2) + 4(2) \end{bmatrix} = \begin{bmatrix} 21 \\ 16 \end{bmatrix}$$

$$x_0 - \eta \nabla f(x_0) = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \eta \begin{bmatrix} 21 \\ 16 \end{bmatrix} = \begin{bmatrix} 2 - 21\eta \\ 2 - 16\eta \end{bmatrix}$$

$$\begin{aligned} f(x_0 - \eta\nabla f(x_0)) &= 4(2-21\eta)^2 + 2(2-16\eta)^2 + 4(2-21\eta)(2-16\eta) - 3(2-21\eta) \\ &= 4(441\eta^2 - 84\eta + 4) + 2(256\eta^2 - 64\eta + 4) + 4(336\eta^2 - 74\eta + 4) - 3(2 - 21\eta) \\ &= 3620\eta^2 - 697\eta + 34 \end{aligned}$$

$$f'(x_0 - \eta\nabla f(x_0)) = 7240\eta - 697 = 0$$

$$\eta = 697/7240 = 0.096$$

PUNTO ESTACIONARIO

$$f''(x_0 - \eta\nabla f(x_0)) = 7240 > 0$$

$\eta = 0.096$ ES UN MÍNIMO

$$x_1 = x_0 - \eta\nabla f(x_0) = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - 0.096 \begin{bmatrix} 21 \\ 16 \end{bmatrix} = \begin{bmatrix} -0.016 \\ 0.464 \end{bmatrix}$$

$k = 1: \quad \nabla f(x_1) = \begin{bmatrix} 8(-0.016) + 4(0.464) - 3 \\ 4(-0.016) + 4(0.464) \end{bmatrix} = \begin{bmatrix} -1.272 \\ 1.792 \end{bmatrix}$

$K = 1:\quad \nabla f(x_1) = \begin{bmatrix} 8(-0.016) + 4(0.464) - 3 \\ 4(-0.016) + 4(0.464) \end{bmatrix} = \begin{bmatrix} -1.272 \\ 1.792 \end{bmatrix}$

$$x_1 - \eta \nabla f(x_1) = \begin{bmatrix} -0.016 \\ 0.464 \end{bmatrix} - \eta \begin{bmatrix} -1.272 \\ 1.792 \end{bmatrix} = \begin{bmatrix} -0.016 + 1.272\eta \\ 0.464 - 1.792\eta \end{bmatrix}$$

$$\begin{aligned} f(x_1 - \eta\nabla f(x_1)) &= 4(-0.016 + 1.272\eta)^2 + 2(0.464 - 1.792\eta)^2 + \\ &\quad 4(-0.016 + 1.272\eta)(0.464 - 1.792\eta) - \\ &\quad 3(-0.016 + 1.272\eta) \\ &= 3.78\eta^2 - 4.83\eta + 0.45 \end{aligned}$$

Realizado con Wolfram

$f'(x_1 - \eta\nabla f(x_1)) = 7.55\eta - 4.83 = 0$ $\quad \eta = 0.64$

$f''(x_1 - \eta\nabla f(x_1)) = 7.55 > 0$ (PD) $\quad$ MÍNIMO LOCAL

$$x_2 = x_1 - \eta\nabla f(x_1) = \begin{bmatrix} -0.016 \\ 0.464 \end{bmatrix} - 0.64 \begin{bmatrix} -1.272 \\ 1.792 \end{bmatrix} = \begin{bmatrix} 0.798 \\ -0.682 \end{bmatrix}$$

$K = 2:\quad \nabla f(x_2) = \begin{bmatrix} 8(0.798) + 4(-0.682) - 3 \\ 4(0.798) + 4(-0.682) \end{bmatrix} = \begin{bmatrix} 0.656 \\ 0.464 \end{bmatrix}$

$$x_2 - \eta\nabla f(x_2) = \begin{bmatrix} 0.798 \\ -0.682 \end{bmatrix} - \eta \begin{bmatrix} 0.656 \\ 0.464 \end{bmatrix} = \begin{bmatrix} 0.798 - 0.656\eta \\ -0.682 - 0.464\eta \end{bmatrix}$$

$f(x_2 - \eta\nabla f(x_2)) = 3.37\eta^2 - 0.65\eta - 1.09$

$f'(x_2 - \eta\nabla f(x_2)) = 6.74\eta - 0.65 = 0$ $\quad \eta = 0.096$

$f''(x_2 - \eta\nabla f(x_2)) = 6.74 > 0$ (PD) $\quad$ MÍNIMO LOCAL

$$x_3 = x_2 - \eta\nabla f(x_2) = \begin{bmatrix} 0.798 \\ -0.682 \end{bmatrix} - 0.096 \begin{bmatrix} 0.656 \\ 0.464 \end{bmatrix} = \begin{bmatrix} 0.735 \\ -0.726 \end{bmatrix}$$

2. Considere la función $f : \mathbb{R}^2 \to \mathbb{R}_\ni$

$$f(x_1, x_2) = 100\,(x_2 - x_1^2)^2 + (1 - x_1)^2.$$

Demuestre que el vector $x^* = (1,1)^T$ es el *mínimo global* de la función f sobre $\mathbb{R}^2$.

$$\begin{aligned} f(x_1, x_2) &= 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\ &= 100(x_1^4 - 2x_1^2 x_2 + x_2^2) + x_1^2 - 2x_1 + 1 \end{aligned}$$

$$\nabla f(x_1, x_2) = \begin{bmatrix} 100(4x_1^3 - 4x_1x_2) + 2x_1 - 2 \\ 100(-2x_1^2 + 2x_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

- $\left.\begin{aligned} 100(4x_1^3 - 4x_1x_2) + 2x_1 - 2 &= 0 \\ 100(-2x_1^2 + 2x_2) &= 0 \end{aligned}\right\} \quad x_1^2 = x_2$
- $100(4x_1^3 - 4x_1^3) + 2x_1 - 2 = 0$
- $2x_1 - 2 = 0$
- $x_1 = 1 \qquad x_2 = 1$

PUNTO ESTACIONARIO: $[1, 1]$

$\nabla^2 f(x_1, x_2) = [\; 100(12x_1^2 - 4x_2) + 2 \qquad 100(-4x_1) \;]$

PUNTO ESTACIONARIO: [1 , 1]

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 100(12x_1^2 - 4x_2) + 2 & 100(-4x_1) \\ 100(-4x_1) & 200 \end{bmatrix}$$

$$\nabla^2 f(1,1) = \begin{bmatrix} 100(12(1)^2 - 4(1)) + 2 & 100(-4(1)) \\ 100(-4(1)) & 200 \end{bmatrix} = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix}$$

CON WOLFRAM

$\lambda_1 = 1001.6$

$\lambda_2 = 0.399$

MÍNIMO LOCAL ESTRICTO

$$Z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$Z^T \nabla^2 f(x_1, x_2) Z = \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 100(12x_1^2 - 4x_2) + 2 & 100(-4x_1) \\ 100(-4x_1) & 200 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$= \begin{bmatrix} 100z_1(12x_1^2 - 4x_2) + 2z_1 - 400x_1 z_2 & -400x_1 z_1 + 200z_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$= 400z_1^2(3x_1^2 - x_2) + 2z_1^2 - 400x_1 z_2 z_1 - 400x_1 z_1 z_2 + 200z_2^2$$

$$= 1200x_1 z_1^2 - 400x_2 z_1^2 + 2z_1^2 - 400x_1 z_2 z_1 - 400x_1 z_1 z_2 + 200z_2^2$$

$$= z_1^2(1200x_1 - 400x_2 + 2) - 800x_1 z_2 z_1 + 200z_2^2$$

$$= 2(z_1^2(600x_1 - 200x_2 + 1) - 400x_1 z_2 z_1 + 100z_2^2)$$

$$= 200(z_1^2(6x_1 - 2x_2 + 0.01) - 4x_1 z_2 z_1 + z_2^2)$$

Parece tener la forma de un trinomio cuadrado perfecto, pero debido a la forma del primer componente, no se puede realizar la descomposición. Debido a esto, se comprobó que el punto proporcionado consiste de un mínimo local estricto, pero debido a que se desconoce si la función es convexa o no, entonces no se pudo llegar a comprobar si el mínimo es global o local.