

1. Gradient Descent (8점)

다변량 함수 $g(x, y) = x^2 + y^2 + xy - 4x - 5y + 11$ 에 대해 아래의 설명을 잘 읽고 단계별로 해를 구하라. 계산과정에서 나오는 소수에 대해서는 소수점 이하 둘째 자리에서 반올림하여 소수점 첫째 자리까지만 표기하라.

1-1. 함수 g 의 편도함수 $\frac{\partial g}{\partial x}$, $\frac{\partial g}{\partial y}$ 를 각각 구하라.

답:

먼저 x 에 대해 편미분을 수행합니다.

$$\begin{aligned}\frac{\partial g}{\partial x} &= \frac{\partial}{\partial x}(x^2 + y^2 + xy - 4x - 5y + 11) \\ &= \frac{\partial}{\partial x}(x^2) + \frac{\partial}{\partial x}(y^2) + \frac{\partial}{\partial x}(xy) - \frac{\partial}{\partial x}(4x) - \frac{\partial}{\partial x}(5y) + \frac{\partial}{\partial x}(11) \\ &= 2x + 0 + y - 4 + 0 + 0 \\ &= 2x + y - 4\end{aligned}$$

다음으로 y 에 대해 편미분을 수행합니다.

$$\begin{aligned}\frac{\partial g}{\partial y} &= \frac{\partial}{\partial y}(x^2 + y^2 + xy - 4x - 5y + 11) \\ &= \frac{\partial}{\partial y}(x^2) + \frac{\partial}{\partial y}(y^2) + \frac{\partial}{\partial y}(xy) - \frac{\partial}{\partial y}(4x) - \frac{\partial}{\partial y}(5y) + \frac{\partial}{\partial y}(11) \\ &= 0 + 2y + x - 0 - 5 + 0 \\ &= 2y + x - 5\end{aligned}$$

1-2. Gradient Descent의 갱신 규칙(업데이트 룰)을 아래와 같이 쓴다. 이를 x 와 y 에 적용한 결과를 각각 제시하라.

$$x_{k+1} = x_k - \alpha \frac{\partial g}{\partial x}$$

$$y_{k+1} = y_k - \alpha \frac{\partial g}{\partial y}$$

답:

위에서 구한 편도함수를 대입하면:

$$x_{k+1} = x_k - \alpha(2x_k + y_k - 4)$$

$$y_{k+1} = y_k - \alpha(2y_k + x_k - 5)$$

1-3. (x, y) 의 초기값 $(x_0, y_0) = (4, 4)$ 로 설정할 때 (x_1, y_1) 을 구하라.

답:

학습률 α 를 그대로 두고 일반식으로 계산합니다.

$$x_1 = x_0 - \alpha(2x_0 + y_0 - 4)$$

$$y_1 = y_0 - \alpha(2y_0 + x_0 - 5)$$

초기값 $(x_0, y_0) = (4, 4)$ 를 대입하면,

$$x_1 = 4 - \alpha(2 \times 4 + 4 - 4) = 4 - 8\alpha$$

$$y_1 = 4 - \alpha(2 \times 4 + 4 - 5) = 4 - 7\alpha$$

따라서 $(x_1, y_1) = (4 - 8\alpha, 4 - 7\alpha)$ 입니다.

1-4. 두 기울기의 크기가 모두 0.2 이하가 될 때까지 (x, y) 의 경로를 수행하고 최종 근사값을 제시하라.

답:

두 기울기의 크기가 모두 0.2 이하가 될 때까지 gradient descent를 수행합니다. 학습률 $\alpha = 0.1$ 로 설정하겠습니다.

초기값: $(x_0, y_0) = (4.00, 4.00)$

1단계:

- $\frac{\partial g}{\partial x}(4.00, 4.00) = 2 \times 4.00 + 4.00 - 4 = 8.00$
- $\frac{\partial g}{\partial y}(4.00, 4.00) = 2 \times 4.00 + 4.00 - 5 = 7.00$
- $x_1 = 4.00 - 0.1 \times 8.00 = 3.20$
- $y_1 = 4.00 - 0.1 \times 7.00 = 3.30$

2단계:

- $\frac{\partial g}{\partial x}(3.20, 3.30) = 2 \times 3.20 + 3.30 - 4 = 6.40 + 3.30 - 4 = 5.70$
- $\frac{\partial g}{\partial y}(3.20, 3.30) = 2 \times 3.30 + 3.20 - 5 = 6.60 + 3.20 - 5 = 4.80$
- $x_2 = 3.20 - 0.1 \times 5.70 = 3.20 - 0.57 = 2.63$
- $y_2 = 3.30 - 0.1 \times 4.80 = 3.30 - 0.48 = 2.82$

3단계:

- $\frac{\partial g}{\partial x}(2.63, 2.82) = 2 \times 2.63 + 2.82 - 4 = 5.26 + 2.82 - 4 = 4.08$
- $\frac{\partial g}{\partial y}(2.63, 2.82) = 2 \times 2.82 + 2.63 - 5 = 5.64 + 2.63 - 5 = 3.27$
- $x_3 = 2.63 - 0.1 \times 4.08 = 2.63 - 0.41 = 2.22$
- $y_3 = 2.82 - 0.1 \times 3.27 = 2.82 - 0.33 = 2.49$

4단계:

- $\frac{\partial g}{\partial x}(2.22, 2.49) = 2 \times 2.22 + 2.49 - 4 = 4.44 + 2.49 - 4 = 2.93$
- $\frac{\partial g}{\partial y}(2.22, 2.49) = 2 \times 2.49 + 2.22 - 5 = 4.98 + 2.22 - 5 = 2.20$
- $x_4 = 2.22 - 0.1 \times 2.93 = 2.22 - 0.29 = 1.93$
- $y_4 = 2.49 - 0.1 \times 2.20 = 2.49 - 0.22 = 2.27$

5단계:

- $\frac{\partial g}{\partial x}(1.93, 2.27) = 2 \times 1.93 + 2.27 - 4 = 3.86 + 2.27 - 4 = 2.13$
- $\frac{\partial g}{\partial y}(1.93, 2.27) = 2 \times 2.27 + 1.93 - 5 = 4.54 + 1.93 - 5 = 1.47$
- $x_5 = 1.93 - 0.1 \times 2.13 = 1.93 - 0.21 = 1.72$
- $y_5 = 2.27 - 0.1 \times 1.47 = 2.27 - 0.15 = 2.12$

6단계:

- $\frac{\partial g}{\partial x}(1.72, 2.12) = 2 \times 1.72 + 2.12 - 4 = 3.44 + 2.12 - 4 = 1.56$
- $\frac{\partial g}{\partial y}(1.72, 2.12) = 2 \times 2.12 + 1.72 - 5 = 4.24 + 1.72 - 5 = 0.96$
- $x_6 = 1.72 - 0.1 \times 1.56 = 1.72 - 0.16 = 1.56$
- $y_6 = 2.12 - 0.1 \times 0.96 = 2.12 - 0.10 = 2.02$

7단계:

- $\frac{\partial g}{\partial x}(1.56, 2.02) = 2 \times 1.56 + 2.02 - 4 = 3.12 + 2.02 - 4 = 1.14$
- $\frac{\partial g}{\partial y}(1.56, 2.02) = 2 \times 2.02 + 1.56 - 5 = 4.04 + 1.56 - 5 = 0.60$
- $x_7 = 1.56 - 0.1 \times 1.14 = 1.56 - 0.11 = 1.45$
- $y_7 = 2.02 - 0.1 \times 0.60 = 2.02 - 0.06 = 1.96$

8단계:

- $\frac{\partial g}{\partial x}(1.45, 1.96) = 2 \times 1.45 + 1.96 - 4 = 2.90 + 1.96 - 4 = 0.86$
- $\frac{\partial g}{\partial y}(1.45, 1.96) = 2 \times 1.96 + 1.45 - 5 = 3.92 + 1.45 - 5 = 0.37$
- $x_8 = 1.45 - 0.1 \times 0.86 = 1.45 - 0.09 = 1.36$
- $y_8 = 1.96 - 0.1 \times 0.37 = 1.96 - 0.04 = 1.92$

9단계:

- $\frac{\partial g}{\partial x}(1.36, 1.92) = 2 \times 1.36 + 1.92 - 4 = 2.72 + 1.92 - 4 = 0.64$
- $\frac{\partial g}{\partial y}(1.36, 1.92) = 2 \times 1.92 + 1.36 - 5 = 3.84 + 1.36 - 5 = 0.20$
- $x_9 = 1.36 - 0.1 \times 0.64 = 1.36 - 0.06 = 1.30$
- $y_9 = 1.92 - 0.1 \times 0.20 = 1.92 - 0.02 = 1.90$

10단계:

- $\frac{\partial g}{\partial x}(1.30, 1.90) = 2 \times 1.30 + 1.90 - 4 = 2.60 + 1.90 - 4 = 0.50$
- $\frac{\partial g}{\partial y}(1.30, 1.90) = 2 \times 1.90 + 1.30 - 5 = 3.80 + 1.30 - 5 = 0.10$
- $x_{10} = 1.30 - 0.1 \times 0.50 = 1.30 - 0.05 = 1.25$
- $y_{10} = 1.90 - 0.1 \times 0.10 = 1.90 - 0.01 = 1.89$

11단계:

- $\frac{\partial g}{\partial x}(1.25, 1.89) = 2 \times 1.25 + 1.89 - 4 = 2.50 + 1.89 - 4 = 0.39$
- $\frac{\partial g}{\partial y}(1.25, 1.89) = 2 \times 1.89 + 1.25 - 5 = 3.78 + 1.25 - 5 = 0.03$
- $x_{11} = 1.25 - 0.1 \times 0.39 = 1.25 - 0.04 = 1.21$
- $y_{11} = 1.89 - 0.1 \times 0.03 = 1.89 - 0.00 = 1.89$

12단계:

- $\frac{\partial g}{\partial x}(1.21, 1.89) = 2 \times 1.21 + 1.89 - 4 = 2.42 + 1.89 - 4 = 0.31$
- $\frac{\partial g}{\partial y}(1.21, 1.89) = 2 \times 1.89 + 1.21 - 5 = 3.78 + 1.21 - 5 = -0.01$
- $x_{12} = 1.21 - 0.1 \times 0.31 = 1.21 - 0.03 = 1.18$
- $y_{12} = 1.89 - 0.1 \times (-0.01) = 1.89 + 0.00 = 1.89$

13단계:

- $\frac{\partial g}{\partial x}(1.18, 1.89) = 2 \times 1.18 + 1.89 - 4 = 2.36 + 1.89 - 4 = 0.25$
- $\frac{\partial g}{\partial y}(1.18, 1.89) = 2 \times 1.89 + 1.18 - 5 = 3.78 + 1.18 - 5 = -0.04$
- $x_{13} = 1.18 - 0.1 \times 0.25 = 1.18 - 0.03 = 1.15$
- $y_{13} = 1.89 - 0.1 \times (-0.04) = 1.89 + 0.00 = 1.89$

14단계:

- $\frac{\partial g}{\partial x}(1.15, 1.89) = 2 \times 1.15 + 1.89 - 4 = 2.30 + 1.89 - 4 = 0.19$
- $\frac{\partial g}{\partial y}(1.15, 1.89) = 2 \times 1.89 + 1.15 - 5 = 3.78 + 1.15 - 5 = -0.07$
- $x_{14} = 1.15 - 0.1 \times 0.19 = 1.15 - 0.02 = 1.13$
- $y_{14} = 1.89 - 0.1 \times (-0.07) = 1.89 + 0.01 = 1.90$

이제 $\frac{\partial g}{\partial x} = 0.19$, $\frac{\partial g}{\partial y} = -0.07$ 로 두 기울기의 크기가 모두 0.2 이하가 되었습니다.

따라서 최종 근사값은 $(x, y) = (1.1, 1.9)$ 입니다.