Suppose $\theta_0 = 1, \theta_1 = 2$, and we simultaneously update θ_0 and θ_1 using the rule: $\theta_j := \theta_j + \sqrt{\theta_0 \theta_1}$ (for j = 0 and j=1) What are the resulting values of θ_0 and θ_1?

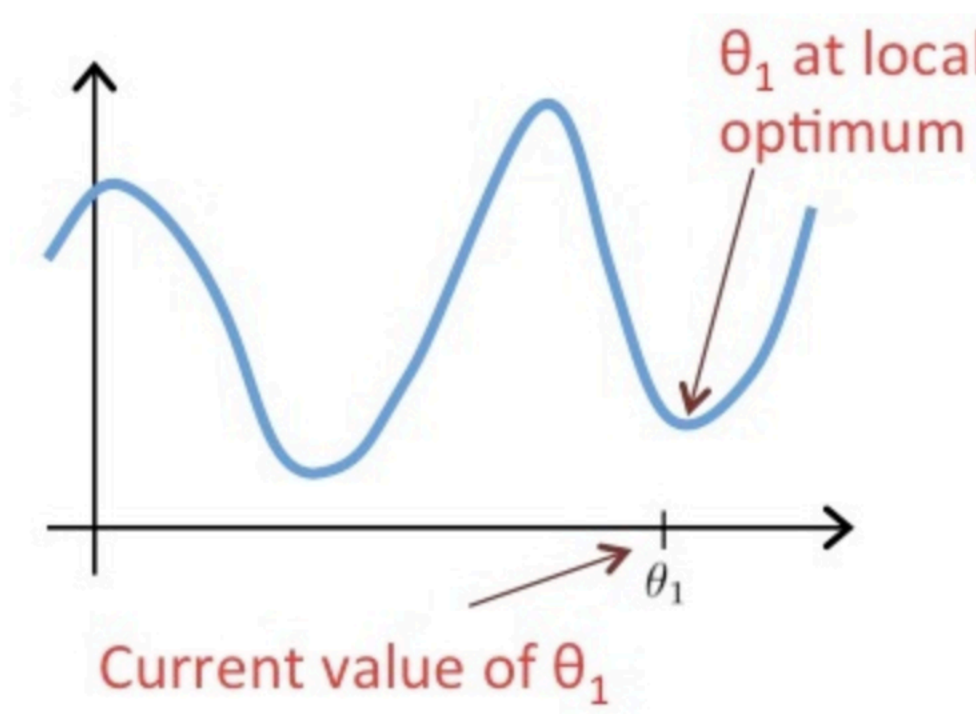
- ○ $\theta_0 = 1, \theta_1 = 2$
- ◉ $\theta_0 = 1 + \sqrt{2}, \theta_1 = 2 + \sqrt{2}$

Correct

- ○ $\theta_0 = 2 + \sqrt{2}, \theta_1 = 1 + \sqrt{2}$
- ○ $\theta_0 = 1 + \sqrt{2}, \theta_1 = 2 + \sqrt{(1 + \sqrt{2}) \cdot 2}$

Suppose θ_1 is at a local optimum of $J(\theta_1)$, such as shown in the figure.

What will one step of gradient descent $\theta_1 := \theta_1 - \alpha \frac{d}{d\theta_1} J(\theta_1)$ do?



- [x] Leave θ_1 unchanged

Correct

Which of the following are true statements? Select all that apply.

- [ ] To make gradient descent converge, we must slowly decrease α over time.

Un-selected is correct

- [ ] Gradient descent is guaranteed to find the global minimum for any function $J(\theta_0, \theta_1)$.

Un-selected is correct

- [x] Gradient descent can converge even if α is kept fixed. (But α cannot be too large, or else it may fail to converge.)

Correct

- [x] For the specific choice of cost function $J(\theta_0, \theta_1)$ used in linear regression, there are no local optima (other than the global optimum).

Correct