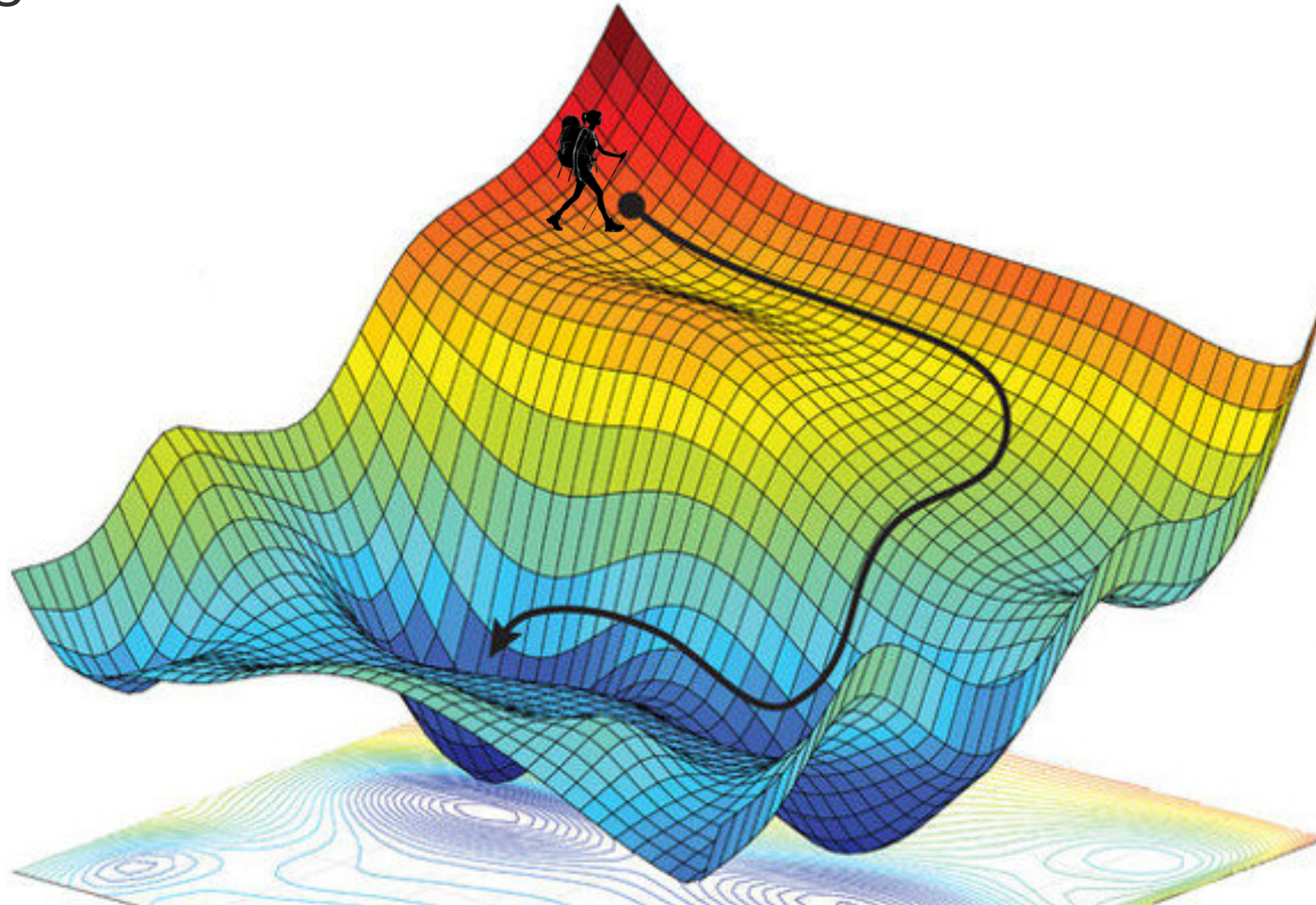


# Estimation: Gradient Descent

- The goal is to minimize the loss function
- The gradient tells us which direction to move



[source: <https://arxiv.org/abs/1805.04829>]



# Gradient Descent

$$f(x, y) = x^2 + xy + y^2$$



Partial Derivatives:

$$\frac{\partial f}{\partial x} = 2x + y \quad \frac{\partial f}{\partial y} = x + 2y$$

$$\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x + y \\ x + 2y \end{bmatrix}$$

