Minimizing functions of multiple variables

- Consider the function:

$$f(x_1, x_2) = (x_1 - 2)^2 + 2x_1 - (x_2 - 3)^2$$

- It has two **partial derivatives**: $\frac{\partial f}{\partial x_1}$ and $\frac{\partial f}{\partial x_2}$.

$$\frac{\partial f}{\partial x_1} = 2(x_1 - 2) + 2 = 2x_1 - 4 + 2 = \boxed{2x_1 - 2}$$

$$\frac{\partial f}{\partial x_2} = \boxed{-2(x_2 - 3)}$$

$$\nabla f(\vec{x}) = \begin{bmatrix} 2x_1 - 2 \\ -2(x_2 - 3) \end{bmatrix}$$

"nabla" ∇

The gradient vector

- If $f(\vec{x})$ is a function of multiple variables, then its **gradient**, $\nabla f(\vec{x})$, is a vector containing its partial derivatives.

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$f : \mathbb{R}^2 \to \mathbb{R}$$

- Example:

$$f(\vec{x}) = (x_1 - 2)^2 + 2x_1 - (x_2 - 3)^2$$

$$\nabla f(\vec{x}) = \begin{bmatrix} 2x_1 - 2 \\ -(2x_2 - 6) \end{bmatrix}$$

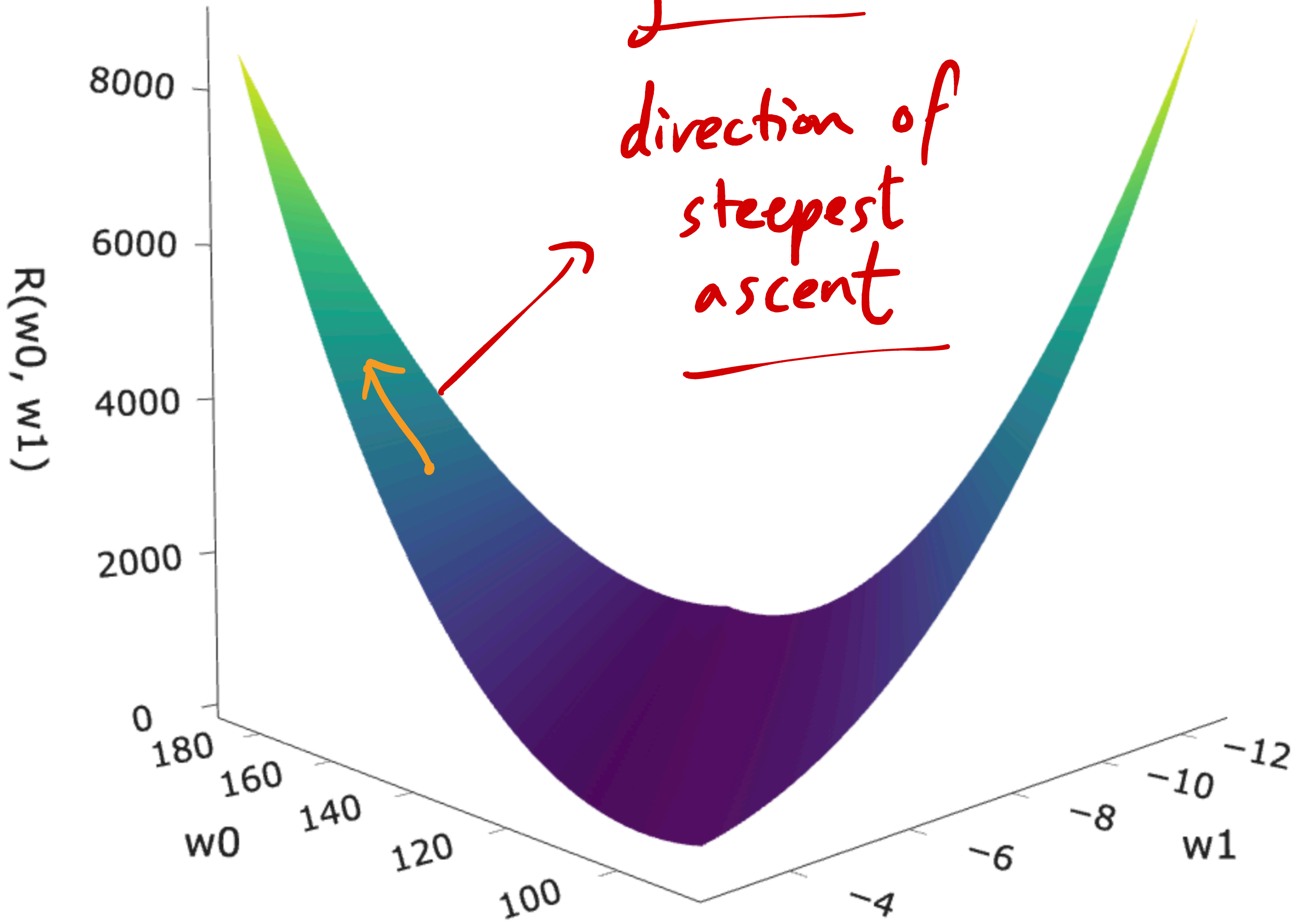
- Example:

scalar!

$$f(\vec{x}) = \vec{x}^T \vec{x} = x_1^2 + x_2^2 + \dots + x_n^2$$

$$\frac{\partial f}{\partial x_i} = 2x_i$$

$$\implies \nabla f(\vec{x}) = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_n \end{bmatrix} = 2\vec{x}$$



gradient:

direction of steepest ascent