Gradient Descent

• An optimization algorithm

• to find the local minimum of a function.

• Take steps proportional to the negative of the gradient of the function at a current point to get to the local minimum.
• Consider a point $a$ where the function $F(x)$ is defined and is differentiable.

• $F(x)$ decreases \textit{fastest} in the direction of the negative gradient of $F$ at $a$.

\[ -\nabla F(a) \]
Gradient Descent

\[ b = a - \eta \nabla F(a) \]

for small enough \( \eta > 0 \),

\[ F(a) \geq F(b) \]
Example

\[ F(x) = 2(x + 4)^2 + 1 \]
Example

\[ F(x) = 2(x + 4)^2 + 1 \]

\[ \frac{df(x)}{dx} = 2x^2 + 16x \]

\[ \frac{df(x)}{dx} = 0 = 2x^2 + 16 \]

\[ x = -4 \]
Example

\[ F(x) = 2(x + 4)^2 + 1 \]
Example

\[ F(x) = 2(x + 4)^2 + 1 \]