Algorithm 1 Gradient Descent

Require: Cost function $J(\mathbf{w})$, initial $\mathbf{w}^{(0)}$, learning rate α, tolerance level ϵ.

1: Set $i \leftarrow 0$
2: **while** $\|\mathbf{w}^{(i)} - \mathbf{w}^{(i-1)}\| > \epsilon$ **do**
3: Update $\mathbf{w}^{(i+1)} \leftarrow \mathbf{w}^{(i)} - \alpha \nabla J(\mathbf{w}^{(i)})$
4: Update $i \leftarrow i + 1$
5: **end while**
6: Return $\widehat{\mathbf{w}} \leftarrow \mathbf{w}^{(i-1)}$