Algorithm 1: The general gradient descent algorithm.

input : initial value $x^{(0)}$, learning rate η, number of iterations T
output: minimizer $x^{(T)}$

for $t = 0$ **to** $T - 1$ **do**
    compute $\Delta x^{(t)} = -\nabla g(x^{(t)})$
    $x^{(t+1)} := x^{(t)} + \eta \Delta x^{(t)}$
end