Week 2 Optimisation for Machine Learning

Neimhin Robinson Gunning, 16321701

February 7, 2024

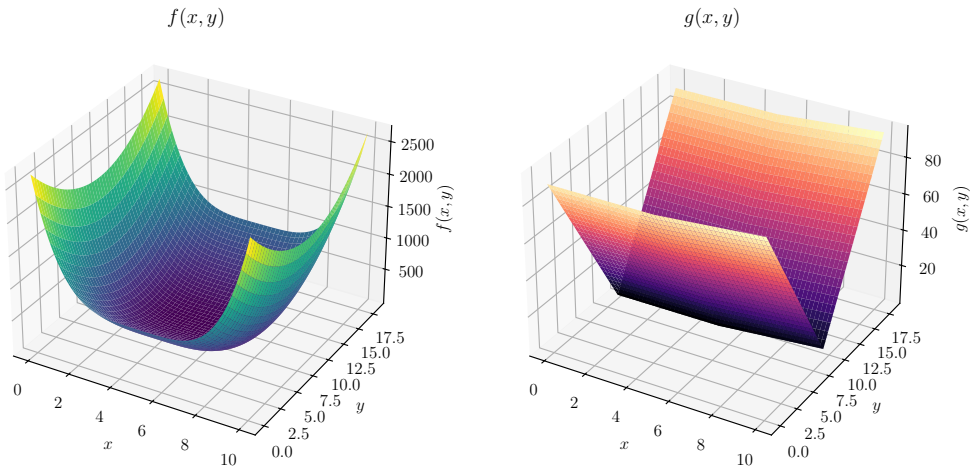

Figure 1

Let

$$f(x, y) = 3(x - 5)^4 + 10(y - 9)^2 \tag{1}$$

and

$$g(x, y) = \max(x - 5, 0) + 10|y - 9| \tag{2}$$

Clearly, the minimum of both $f(x, y)$ and $g(x, y)$ is 0 and they both minimized by $x = 5$, $y = 9$.

The Polyak step size is

$$\alpha_{\text{Polyak}} = \frac{f(x) - f^*}{\nabla f(x)^T \nabla f(x)} \tag{3}$$

where x is the parameter vector, $f(x)$ is the function to optimise, and $f^* \approx \min_x f(x)$.

Listing 1: A python function to calculate the Polyak step size on a `sympy` function.

src/polyak_step_size.py Wed Feb 07 16:44:22 2024 1

```
 1: import numpy as np
 2:
 3:
 4: def polyak_step_size(sp_func, sp_x, x, f_star):
 5:     assert len(sp_x) == len(x)
 6:     subs = {sp_xi: xi for sp_xi, xi in zip(sp_x, x)}
 7:     fx = sp_func.subs(subs)
 8:     grad = [sp_func.diff(sp_xi).subs(subs) for sp_xi in sp_x]
 9:     grad = np.array(grad)
10:     denominator = sum(grad * grad)
11:     numerator = fx - f_star
12:     return numerator / denominator
```