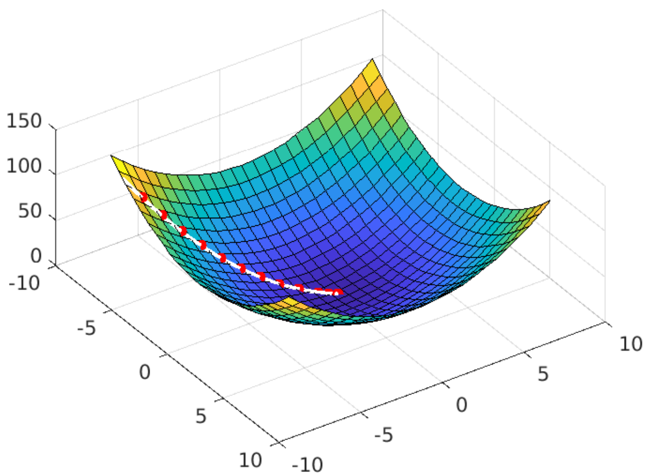
Question 9

The gradient descend function is used to check the performance of the Wolfe condition function. As shown in the screenshot, the wolf condition, alongside with the gradient descent , finds the minimum of the function.



For this exercise, the code of Armijo_LS.m is modified in a way that it satisfies the Wolfe condition.

Note: in Armijo_LS.m , the multiplier “rho” is replaced by c2 like the formula provided in the slides:

Stronge Wolfe Conditions

$$f(x+\alpha p) \leq f(x) + c_1 \alpha p^T \nabla f(x)$$
$$|p^T \nabla f(x+\alpha p)| \leq c_2 |p^T \nabla f(x)|$$

Note: in the second Wolfe condition, we should check and make sure that the derivative in the accepted point is not too big, as it guarantees that we are terminating near a stationary point where derivative is zero. So, we replace the second condition of Wolfe conditions to the strong Wolfe condition to make sure we are terminating at a proper point.

Stronge Wolfe Conditions

$$f(x+\alpha p) \leq f(x) + c_1 \alpha p^T \nabla f(x)$$

$$|p^T \nabla f(x+\alpha p)| \leq c_2 |p^T \nabla f(x)|$$

with constants $0 < c_1 < c_2 < 1$.The only difference with the Wolfe conditions is that we no longer allow the derivative $|p^T \nabla f(x+\alpha p)|$ to be too positive.

The function is very similar to Armijo , except the fact at each iteration we check if the second strong Wolfe condition is satisfied.