Question-3 AI21BTECH11011 Gunjit Mittal

Question

Consider the following unconstrained optimization problem.

$$\min \quad \tfrac{1}{3}x_1^6 - 2.1x_1^4 + 4x_2^4 + 4(x_1^2 - x_2^2) + x_1x_2$$

$$\text{Sub. to } \; x_1 + x_2 = 0.6228$$

Using a programming language of your choice, obtain the minima of the optimization problem. If necessary, use the initial point as $x^0 = [1, -1]^T$.

Answer

Since the problem is an equality-constrained optimization problem, we will use the Lagrangian condition to solve it. We are given the following:

$$f(x) = \frac{1}{3}x_1^6 - 2.1x_1^4 + 4x_2^4 + 4(x_1^2 - x_2^2) + x_1x_2 \text{ and } h(x) = x_1 + x_2 - 0.6228$$

and an initial point $x^0 = [1, -1]^T$.

The Lagrangian function $l(x, \lambda) = f(x) + \lambda^T h(x)$ and the gradient of the lagrangian function is given by $\nabla l(x, \lambda) = \begin{bmatrix} \nabla f(x) + Dh(x)^T \lambda \\ h(x) \end{bmatrix}$.

We have $\nabla f(x) = \begin{bmatrix} 2x_1^5 - 8.4x_1^3 + 8x_1 + x_2 \\ 16x_2^3) - 8x_2 + x_1 \end{bmatrix}$ and $\nabla h(x) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \implies$ $Dh(x) = [1, 1]$

Using this, we get the gradient of the Lagrangian function as

$$\nabla l(x, \lambda) = \begin{bmatrix} 2x_1^5 - 8.4x_1^3 + 8x_1 + x_2 + \lambda \\ 16x_2^3) - 8x_2 + x_1 + \lambda \\ x_1 + x_2 - 0.6228 \end{bmatrix}$$

We know that the solution to this optimization problem can be found by solving for the roots of $\nabla l(x, \lambda)$. So applying Newton's method to find roots for it, we can find the solution for it.

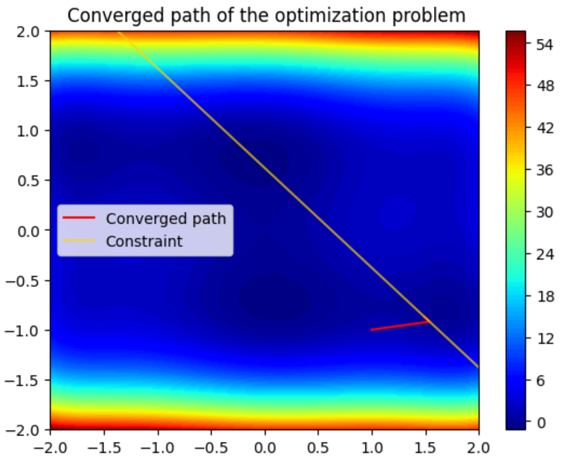
We find the jacobian of the lagrangian function to apply Newton's method.

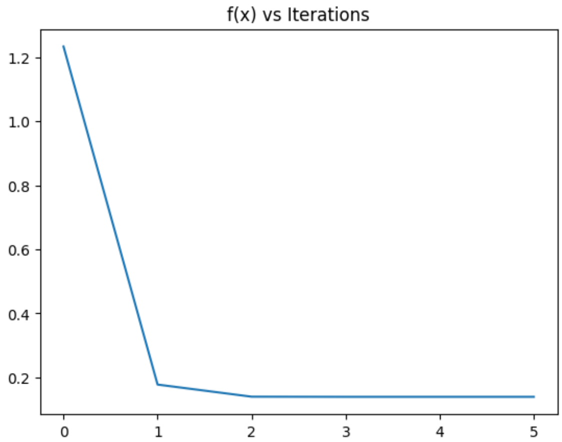
$$Dl(x,\lambda) = \begin{bmatrix} 10x_1^4 - 25.2x_1^2 + 8 & 1 & 1 \\ 1 & 48x_2^2 - 8 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$ and use it to apply Newton's method.

Using the initial conditions given and assuming any starting λ (I used π) and a tolerance value ($\epsilon = 1e^{-5}$ in my case) we run Newton's method with the following update equation:

$$\text{solve for } \Delta x \text{ where: } [Dl(x^k,\lambda)]^T[Dl(x^k,\lambda)]\Delta x = -[Dl(x^k,\lambda)]^T \nabla l(x,\lambda)$$
$$x^{k+1} = x^k + \Delta x$$

Solving through code, we get the following plots:





The minima is $x = [1.491415 - 0.868615]$ with the minimum value being 0.139269, which can be seen from the plots as well. The value of λ obtained is 2.045475. The code converges in 5 iterations. We can see from the contour plot that even though the starting point is not in the feasible region, it moves to a point within the feasible range in the 1st iteration itself. After this, it continues minimising the function within the feasible region. Since the function has more than one local minimum as seen from the contour plot, we can argue that the function $f(x)$ is not convex.

Note: If we had used a different starting point, i.e. $x^0 = [-1, 1]^T$ we could have reached a better minimum of -1.031628 because Newton's method tends to approach the closest local minima.