Home Problem 1
FFR105

Eric Blohm
20000511-3030

September 23, 2024

Problem 1.1

1.

The function $f_p(\mathbf{x};\mu)$ consists of both $f(\mathbf{x})$ and the penalty function $p(\mathbf{x};\mu)$, where:

$$\begin{aligned} p(\mathbf{x};\mu) &= max(g(\mathbf{x}),0)^2 \\ &= max((x_1^2 + x_2^2 - 1), 0)^2 \end{aligned}$$

Thus:

$$\begin{aligned} f_p(\mathbf{x};\mu) &= f(\mathbf{x}) + p(\mathbf{x};\mu) \\ &= \begin{cases} (x_1-1)^2 + 2(x_2-2)^2 + \mu(x_1^2 + x_2^2 - 1)^2 & \text{if } x_1^2 + x_2^2 \geq 1 \\ (x_1-1)^2 + 2(x_2-2)^2 & \text{otherwise} \end{cases} \end{aligned}$$

2.

Two cases of the gradient occur, one where the constraint is fulfilled and therefore the penalty function will be zero, and thus $f_p(\mathbf{x}) = f(\mathbf{x})$. The other case, where the constraint is not fulfilled, the penalty function will be included.
The two cases are the following:

- Gradient of the unconstrained function:

$$\nabla f_p(\mathbf{x};\mu) = \begin{bmatrix} 2(x_1-1) + 4\mu x_1(x_1^2 + x_2^2 - 1) \\ 4(x_2-2) + 4\mu x_2(x_1^2 + x_2^2 - 1) \end{bmatrix}$$

- Gradient of the constrained function:

$$\nabla f_p(\mathbf{x};\mu) = \begin{bmatrix} 2(x_1-1) \\ 4(x_2-2) \end{bmatrix}$$

3.

To find the unconstrained minimum ($\mu = 0$) for the function. I first check for stationary points by putting $\nabla f_p(\mathbf{x};\mu) = \mathbf{0}$. That is:

$$\begin{aligned} \nabla f_p(\mathbf{x};\mu) = \begin{bmatrix} 2(x_1-1) \\ 4(x_2-2) \end{bmatrix} &= \mathbf{0} \\ \implies \begin{cases} x_1 = 1 \\ x_2 = 2 \end{cases} \end{aligned}$$

After the stationary points are located, I will use the Hessian to determine if it is positive definite, i.e has a local minimum at the stationary point. Derivation of $\nabla f_p(\mathbf{x};\mu)$ gives the following Hessian (H).

$$H = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

To check whether the Hessian is positive definite or not, I examine the eigenvalues using $det(H - \lambda I) = 0$:

$$\begin{aligned} det(H - \lambda I) &= (2 - \lambda)(4 - \lambda) \\ &= \lambda^2 - 6\lambda + 8 = 0 \\ &\implies \\ \lambda_{1,2} &= 3 \pm \sqrt{9 - 8} = 3 \pm 1 > 0 \end{aligned}$$

Since $\lambda_{1,2} > 0$, the Hessian is positive definite and thus the stationary point $(x_1, x_2) = (1, 2)$ is a local minimum.

4.

See attached zip with matlab code.

5.

When running the penalty method with the values of μ, η and T, listed in the problem description, with the penalty function:

$$p(\mathbf{x};\mu) = (x_1-1)^2 + 2(x_2-2)^2 + \mu(x_1^2+x_2^2-1)^2$$

I get the results listed in Table 1. As μ increases, the change in x_1 and x_2 decreases significantly for big μ, indicating that they seem to converge. Additionally, from studying Figure 1, the values of x_1 and x_2 seem to converge.

μ	x_1	x_2
1	0.4338	1.2102
10	0.3314	0.9955
100	0.3137	0.9553
1000	0.3118	0.9507

Table 1: The left-most column is a positive parameter used in the penalty function, increasing from 1 to 1000 with a factor of 10. The middle and left-most columns represent the change in x_1 and x_2 respectively, as μ increases.

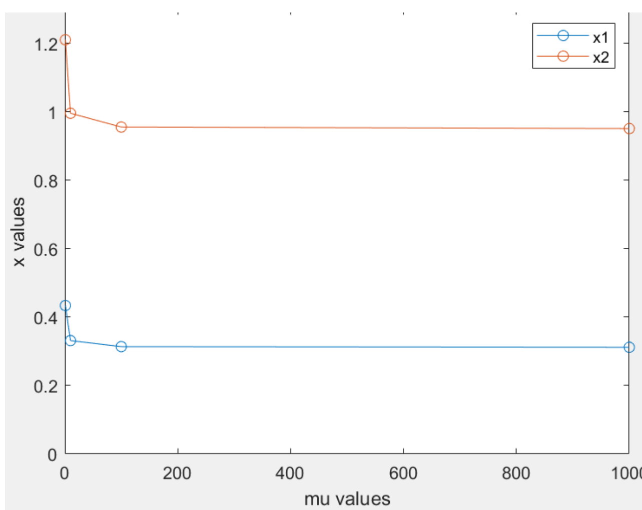

Figure 1: A visual representation of the values in Table 1.

Problem 1.2

Let $f(x_1, x_2) = 4x_1^2+2x_2^3$ and $g(x_1, x_2) = x_1^2+x_2^2-4 \leq 0$. To find the stationary point, check the points when the gradient of f is zero. That is:

$$\nabla f(x_1, x_2) = \begin{bmatrix} 8x_1 \\ 6x_2^2 \end{bmatrix} = \mathbf{0}$$
$$\implies$$
$$\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

Which fulfills g. Now, inserting x_1 and x_2 in f gives $f(0, 0) = 0$.
When examining the boundary of S, I will use the Lagrange multiplier method with the equality constraint $h(x_1, x_2) = x_1^2 + x_2^2 - 4 = 0$:

$$\begin{aligned} L(x_1, x_2, \lambda) &= f(x_1, x_2) + \lambda h(x_1, x_2) \\ &= 4x_1^2 + 2x_2^3 + \lambda(x_1^2 + x_2^2 - 4) \end{aligned}$$

Now find the stationary points of L:

$$\nabla L(x_1, x_2, \lambda) = \begin{bmatrix} 8x_1 + 2\lambda x_1 \\ 6x_2^2 + 2\lambda x_2 \\ x_1^2 + x_2^2 - 4 \end{bmatrix} = \mathbf{0}$$

I can immediately see 2 trivial cases, one where $x_1 = 0$ and one where $x_2 = 0$. Examining these 2 cases one by one gives us:

- Case 1, $x_1 = 0$:
  From the equality constraint we get x_2: $x_2^2 = 4 \implies x_2 = \pm 2$.
  - Let $x_2 = 2$:
    $f(0, 2) = 2 * 2^3 = 16$
  - Let $x_2 = -2$:
    $f(0, -2) = 2 * (-2)^3 = -16$
- Case 2: $x_2 = 0$
  Analogously we get x_1 from the equality constraint: $x_1^2 = 4 \implies x_1 = \pm 2$.
  - Let $x_1 = 2$:
    $f(2, 0) = 4 * 2^2 = 16$
  - Let $x_1 = -2$:
    $f(-2, 0) = 4 * (-2)^2 = 16$

Now proceed with the cases where x_1 and x_2 are not zero:

$$\begin{cases} 8x_1 + 2\lambda x_1 = 0 \\ 6x_2^2 + 2\lambda x_2 = 0 \\ x_1^2 + x_2^2 - 4 = 0 \end{cases} \implies \begin{cases} \lambda = -4 \\ x_2(6x_2 - 8) = 0 \\ x_1^2 + x_2^2 - 4 = 0 \end{cases}$$

$$\implies \begin{cases} x_1 = \pm\sqrt{20}/3 \\ x_2 = 4/3 \end{cases}$$

I have found two more cases from inserting $x_2 = 4/3$ in to the equality constraint, $x_1 = \pm\sqrt{20}/3$, however it will result in the same function value due to the square of x_1 in f. That is:

$$f(\pm\frac{\sqrt{20}}{3}, \frac{4}{3}) = 4 * (\pm\frac{\sqrt{20}}{3})^2 + 2(\frac{4}{3})^3 = \frac{368}{27} \approx 13.6$$

Summarizing all points and corresponding function values:

- $f(0,0) = 0$
- $f(0,2) = 16$
- $f(0,-2) = -16$
- $f(2,0) = 16$
- $f(-2,0) = 16$
- $f(\pm\frac{\sqrt{20}}{3}, \frac{4}{3}) \approx 13.6$

The maximum on the closed set S are at the points: $(0,2), (2,0), (-2,0)$ with the value 16. The minimum is at $(0,-2)$ with a value of -16.

Problem 1.3

a)

The function used is:

$$\begin{aligned} g(x_1, x_2) = &(1.5 - x_1 + x_1 x_2)^2 \\ &+ (2.25 - x_1 + x_1 x_2^2)^2 \\ &+ (2.625 - x_1 + x_1 x_2^3)^2 \end{aligned}$$

The parameters chosen are the pre-set ones. That is:

- tournamentSize = 2;
- tournamentProbability = 0.75;
- crossoverProbability = 0.8;
- mutationProbability = 0.02;
- numberOfGenerations = 2000;

Additionally, for the parameters where no changes was allowed:

- populationSize = 100;
- maximumVariableValue = 5;
- numberOfGenes = 50;
- numberOfVariables = 2;

The values of $g(x_1, x_2)$, x_1, and x_2 when running the program RunSingle.m are stated in Table 2, where a row represents a single run.

run #	x_1	x_2	$g(x_1, x_2)$
1	2.9999744296	0.4999904484	$3.424276 * 10^{-10}$
2	2.9999541640	0.4999886602	$3.363110 * 10^{-10}$
3	3.0000498295	0.5000125021	$3.978990 * 10^{-10}$
4	3.0000307560	0.5000077337	$1.516671 * 10^{-10}$
5	2.9999961853	0.4999990910	$2.360778 * 10^{-12}$
6	2.9999979734	0.4999993891	$9.285905 * 10^{-13}$
7	3.0000122786	0.5000029653	$2.427614 * 10^{-11}$
8	3.0000072122	0.5000020713	$1.017120 * 10^{-11}$
9	3.0000116825	0.5000032634	$2.494271 * 10^{-11}$
10	3.0000021458	0.5000005811	$7.924772 * 10^{-13}$

Table 2: Values of $g(x_1, x_2)$, x_1, and x_2 after running the RunSingle.m program 10 times.

b)

The results after running RunBatch.m with 10 different values of p_{mut} are stated in Table 3. When studying the plot in Figure 2 it is evident that a p_{mut} value of 0 is not optimal in this case. Moreover, when examining Figure 3, it seems that $p_{mut} = 1/numberOfGenes$ is the optimal value, which is expected. For $p_{mut} > 1/numberOfGenes$ the median performance tends to decrease as p_{mut} increase. This can also be clarified when studying Table 3. Note also that the optimal value of p_{mut} is usually when $p_{mut} = 1/numberOfGenes$, or a few times this value. That is why not larger values of p_{mut} were tested.

p_{mut}	median performance
0	0.992176763006216
0.02	0.999999981681416
0.04	0.999998621091811
0.06	0.999988207577258
0.08	0.999974714344156
0.10	0.999961477943751
0.12	0.999835493405400
0.14	0.999785829537472
0.16	0.999709767599410
0.18	0.999775100889433

Table 3: Fitness values (median performance) as a function of p_{mut} after running the file RunBatch.m with 10 different values of p_{mut}.

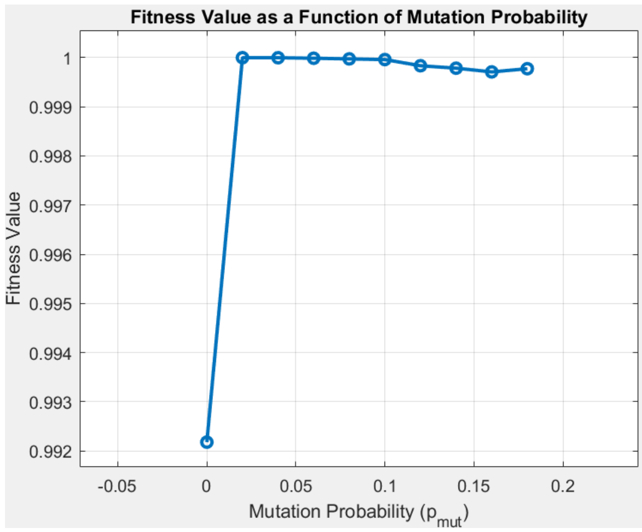

Figure 2: Plot of the values in Table 3.

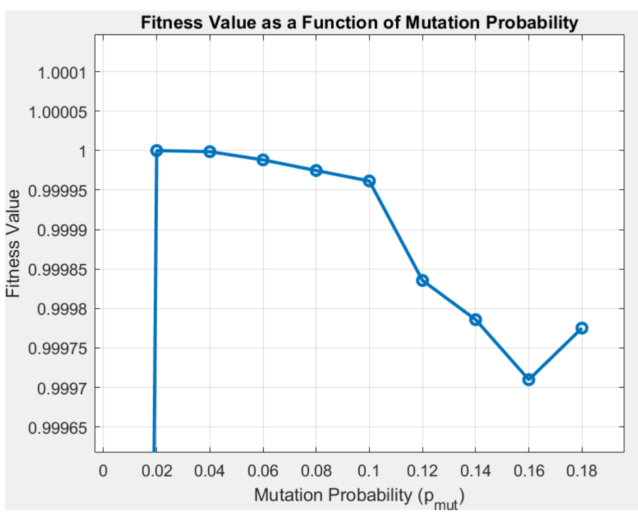

Figure 3: Zoomed-in version of Figure 2 for clarification.

c)

From the results in part a), the best value for x_1, and x_2 seem to be the point $(3, 0.5)$. Which is what I will insert to $\nabla g(x_1, x_2)$, and check whether it is zero:

$$\begin{aligned}\frac{\partial g}{\partial x_1} &= 2(1.5 - x_1 + x_1x_2)(-1 + x_2)\\ &+ 2(2.25 - x_1 + x_1x_2^2)(-1 + x_2^2)\\ &+ 2(2.625 - x_1 + x_1x_2^3)(-1 + x_2^3)\end{aligned}$$

$$\begin{aligned}\frac{\partial g}{\partial x_2} &= 2(1.5 - x_1 + x_1x_2)(x_1)\\ &+ 2(2.25 - x_1 + x_1x_2^2)(2x_1x_2)\\ &+ 2(2.625 - x_1 + x_1x_2^3)(3x_1x_2^2)\end{aligned}$$

Now, when inserting $x_1 = 3$ and $x_2 = 0.5$:

$$\begin{aligned}\frac{\partial g(3, 0.5)}{\partial x_1} &= 2\underbrace{(1.5 - 3 + \underbrace{3 * 0.5}_{=1.5})}_{=0}(-1 + 0.5)\\ &+ 2\underbrace{(2.25 - 3 + \underbrace{3 * 0.25}_{=0.75})}_{=0}(-1 + 0.25)\\ &+ 2\underbrace{(2.625 - 3 + \underbrace{3 * 0.125}_{=0.375})}_{=0}(-1 + 0.125)\\ &= 0\end{aligned}$$

$$\begin{aligned}\frac{\partial g(3, 0.5)}{\partial x_2} &= 2\underbrace{(1.5 - 3 + \underbrace{3 * 0.5}_{=1.5})}_{=0}(3)\\ &+ 2\underbrace{(2.25 - 3 + \underbrace{3 * 0.25}_{=0.75})}_{=0}(6 * 0.5)\\ &+ 2\underbrace{(2.625 - 3 + \underbrace{3 * 0.125}_{=0.375})}_{=0}(9 * 0.25)\\ &= 0\end{aligned}$$

Since $\nabla g(3, 0.5) = \mathbf{0}$, then $(3, 0.5)$ is a stationary point of $g(x_1, x_2)$.