Home Problem 1

Isac Nordin, CID: Isacn
isacn@student.chalmers.se

2021–09–21

1 Problem 1.1: Penalty Method

Problem: Minimize function f with constraint g:

$$f(x_1, x_2) = (x_1 - 1)^2 + 2(x_2 - 2)^2$$

$$g(x_1, x_2) = x_1^2 + x_2^2 - 1 \leq 0$$

Define p,f_p:

$$p(x_1, x_2, \mu) = \mu \cdot (\max(0, x_1^2 + x_2^2 - 1))^2$$

$$f_p(\boldsymbol{x}, \mu) = (x_1 - 1)^2 + 2(x_2 - 2)^2 + \begin{cases} \mu(x_1^2 + x_2^2 - 1)^2, & \text{for } (x_1^2 + x_2^2) \geq 1 \\ 0, & \text{otherwise} \end{cases}$$

Gradiant ∇f_p:

$$\nabla f_p(\boldsymbol{x}, \mu) = [2(x_1 - 1), 4(x_2 - 2)] + \begin{cases} [4\mu x_1(x_1^2 + x_2^2 - 1), \quad 4\mu x_2(x_1^2 + x_2^2 - 1), & \text{for } (x_1^2 + x_2^2) \geq 1 \\ [0, 0], & \text{otherwise} \end{cases}$$

Unconstrained solution (start value for gradient descent):

$$\nabla f_p(\boldsymbol{x}, 0) = \mathbf{0} = \begin{cases} 2(x_1 - 1) = 0 \\ 4(x_2 - 2) = 0 \end{cases} \longrightarrow \begin{cases} x_1 = 1 \\ x_2 = 2 \end{cases}$$

Running gradiant descent with different μ values($\boldsymbol{USED} \quad \eta = 0.00001$):

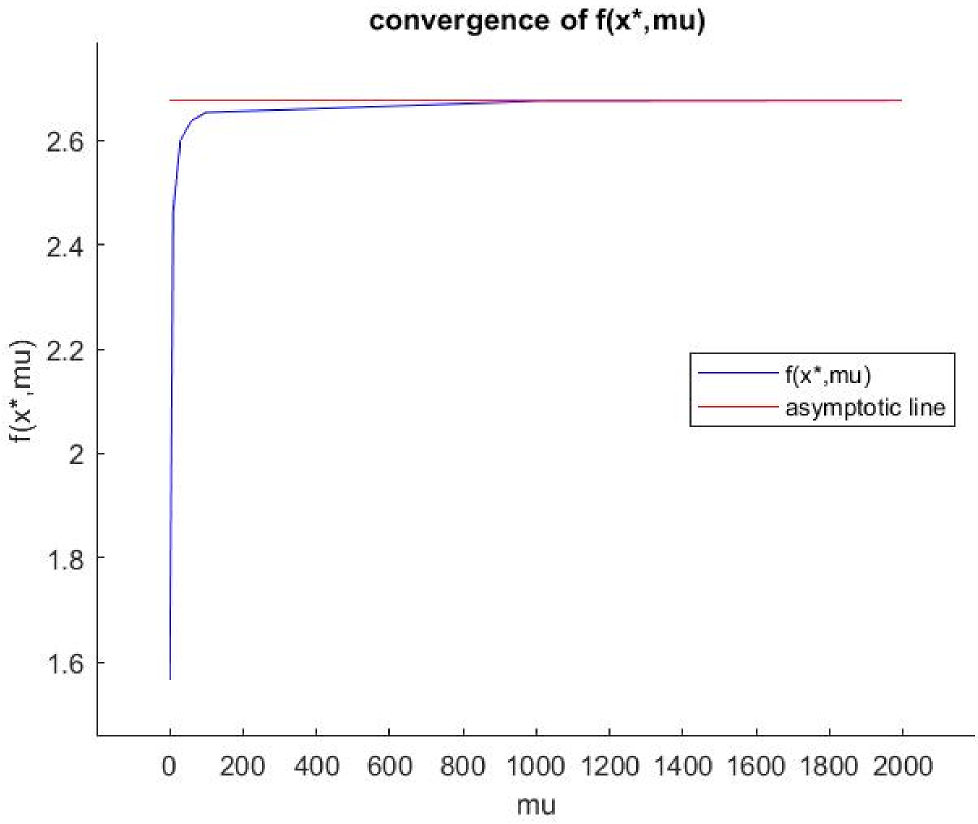
we see $f(\boldsymbol{x}^*, \mu)$ in both table and figure below converges when $\mu \to \infty$. We get $f(\boldsymbol{x}^*, \mu) \approx 2.6768$ for $\boldsymbol{x}^* = (0.3117, 0.9505).\mu = 2000$.

1.1 brief discussion

We can see that $f(\boldsymbol{x}^*, \mu)$ is less when μ is small so why do we not simply choose μ large in the first place? If μ is large it's the same as putting a big penalty for being outside the constraint making it harder to find the minimum we are looking for outside the constraint. Here even μ=2000 doesn't follow the constraint (g(x_1, x_2)<0) and a larger μ-value would be needed. Another reason is that there's not a specific value of μ which is considered large and varies problem to problem. By plotting $f(\boldsymbol{x}^*, \mu)$ for different μ-values we can guarantee if μ is big enough if the $f(\boldsymbol{x}^*, \mu)$ converges to a specific value (as in figure 1).
***Note* :** Numerical calculation with μ»2000, η might need to be readjusted

Tabell 1: $f(\boldsymbol{x}^*, \mu)$ for different $\boldsymbol{x}^*, \mu$, and if it follows constraint

μ	$\boldsymbol{x_1^*}$	$\boldsymbol{x_2^*}$	$f(\boldsymbol{x}^*, \mu)$	$g(\boldsymbol{x}^*)$
1.0000	0.4338	1.2102	1.5683	0.6527
10.0000	0.3314	0.9955	2.4650	0.1009
30.0000	0.3186	0.9665	2.6005	0.0356
60.0000	0.3152	0.9585	2.6383	0.0181
100.0000	0.3137	0.9553	2.6540	0.0109
1000.0000	0.3118	0.9507	2.6756	0.0011
2000.0000	0.3117	0.9505	2.6768	0.0006



Figur 1: Convergence of f($\boldsymbol{x^*}, \mu$) alongside an asymptotic line

2 Problem 1.2: Constrained Optimization

2.1 A) an analytical method

Find global minimum of f(x_1, x_2)=$4x_1^2 - x_1x_2 + 4x_2^2 - 6x_2$ in closed set S that's bound by a triangle with corners: (0,0), (0,1), (1,1).

Stationary points $\in$ interior(S):

$$\nabla f(x_1, x_2) = \mathbf{0} \longrightarrow \begin{cases} 8x_1 - x_2 = 0 & ① \\ -x_1 + 8x_2 - 6 = 0 & ② \end{cases} \longrightarrow \begin{cases} 8① - ② \\ 8② - ① \end{cases} \longrightarrow \begin{cases} x_1 = 6/63 = 2/21 \\ x_2 = 6 \cdot 8/63 = 16/21 \end{cases}$$

Check: $S_1 \in S$?Yes it is(since 0<x_1, x_2<1 and $x_2 > x_1$).
Stationary points $\in$ boundary of S:

$$\begin{cases} x_1 = 0 \\ 0 < x_2 < 1 \end{cases} \longrightarrow \frac{df(0, x_2)}{dx_2} = 8x_2 - 6 = 0 \longrightarrow \text{stationary point } S_2 = (0, \frac{3}{4})$$

$$\begin{cases} 0 < x_1 < 1 \\ x_2 = 1 \end{cases} \longrightarrow \frac{df(x_1, 1)}{dx_1} = 8x_1 - 1 = 0 \longrightarrow \text{stationary point } S_3 = (\frac{1}{8}, 1)$$

$$0 < x_1 = x_2 < 1 \longrightarrow \frac{df(x_1, x_1)}{dx_1} = 14x_1 - 6 = 0 \longrightarrow \text{stationary point } S_4 = (\frac{3}{7}, \frac{3}{7})$$

Stationary points $\in$ corners of triangle: $S_5 = (0, 0), S_6 = (0, 1), S_7 = (1, 1)$.
Evaluate all stationary points:

Tabell 2: Evaluation of stationary points

Stationary points S_x	f(S_x)
S_1=(2/21,16/21)	-2.2857
S_2=(0,3/4)	-2.2500
S_3=(1/8,1)	-2.0625
S_4=(3/7,3/7)	-1.2857
S_5=(0,0)	0
S_6=(0,1)	-2
S_7=(1,1)	1

Global minimum to $f(x_1, x_2)$ is $S_1 = (2/21, 16/21)$, f(S_2)=-2.2857

2.2 B) Lagrange multiplier method

Find global minimum of f(x1,x2)=$15 + 2x_1 + 3x_2$ subject to constraint h$(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 - 21 = 0$

Lagrange:

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2)$$

gradiant of lagrange equals 0:

$$\nabla L(x_1, x_2, \lambda) = \begin{cases} \frac{\partial f}{\partial x_1} + \lambda \frac{\partial h}{\partial x_1} \\ \\ \frac{\partial f}{\partial x_2} + \lambda \frac{\partial h}{\partial x_2} \longrightarrow \\ \\ \frac{\partial f}{\partial \lambda} = h \end{cases} \begin{cases} 2 + \lambda(2x_1 + x_2) = 0 & ① \\ \\ 3 + \lambda(x_1 + 2x_2) = 0 & ② \\ \\ x_1^2 + x_1x_2 + x_2^2 - 21 = 0 & ③ \end{cases}$$

Applying: $2 \cdot ① - ②, \quad 2 \cdot ② - ①$

$$\nabla L(x_1, x_2, \lambda) = \begin{cases} 1 + 3\lambda x_1 = 0 \\ 4 + 3\lambda x_2 = 0 \\ x_1^2 + x_1 x_2 + x_2^2 - 21 = 0 \end{cases} \longrightarrow \begin{cases} x_1 = -\frac{1}{3\lambda} \\ \\ x_2 = -\frac{4}{3\lambda} = 4x_1 \\ \\ x_1^2 + x_1(4x_1) + (4x_1)^2 - 21 = 21x_1^2 - 21 = 0 -> x_1 = \pm 1 \end{cases}$$

Tabell 3: Solution to $\nabla L(x_1, x_2, \lambda) = \mathbf{0}$

x_1	x_2	λ	$f(x_1, x_2)$
1	4	-1/3	29
-1	-4	+1/3	1

The global minimum for f with constraint h is $(x_1, x_2) = (-1, -4)$

3 Problem 1.3: Basic GA program

3.1 A)

want to find minimum of $g(x_1, x_2) = (1.5 - x_1 + x_1x_2^1)^2 + (2.25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2$

Tabell 4: Parameters used

tournamentSize	2
tournamentProbability	0.75
crossoverProbability	0.8
mutationProbability	0.02
numberOfGenerations	30

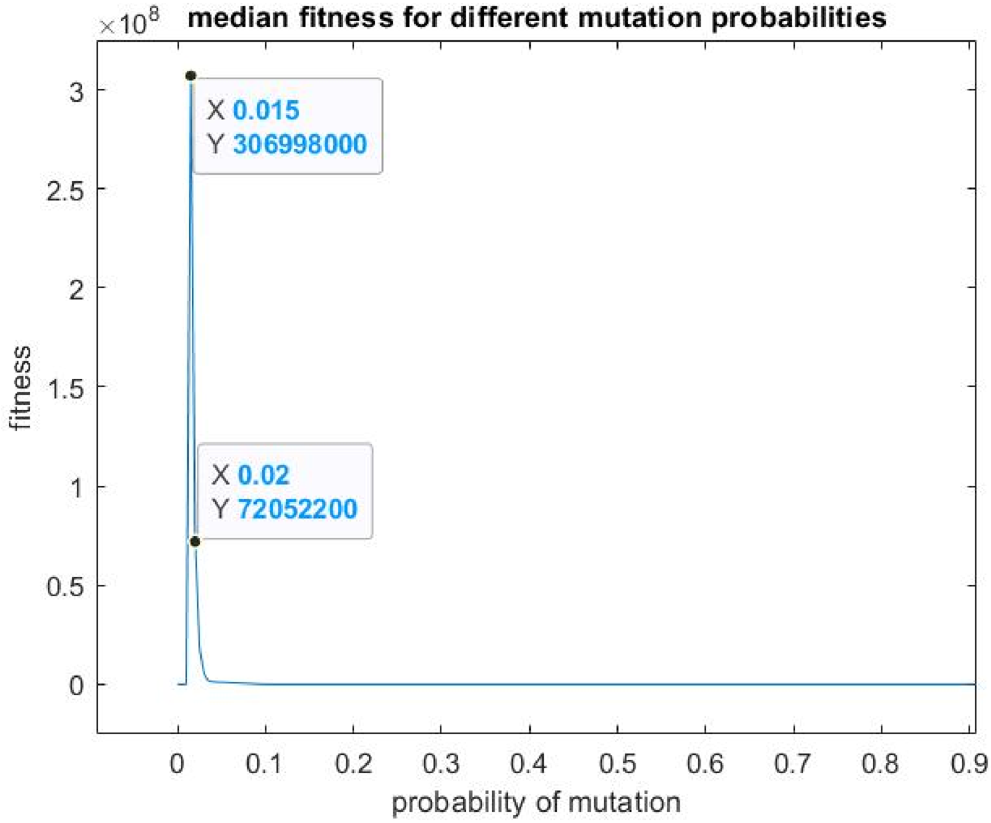
Tabell 5: Minimum according to GA

Run	x_1	x_2	$g(x_1, x_2)$
1	3.3870	0.5855	0.0166
2	3.0898	0.5226	0.0012
3	3.1839	0.5419	0.0044
4	2.9417	0.4848	0.0006
5	3.0482	0.5115	0.0004
6	3.0047	0.5009	0.0000
7	2.9579	0.4896	0.0003
8	2.9902	0.4983	0.0000
9	3.1078	0.5274	0.0017
10	3.0665	0.5135	0.0008

3.2 B)

Tabell 6: median fitness for different $P_{mutation}$

μ	median fitness/10^8
0	0.0000
0.0050	0.0000
0.0100	0.0001
0.0150	3.0700
0.0200	0.7205
0.0250	0.1829
0.0300	0.0551
0.0350	0.0162
0.0400	0.0123
0.1000	0.0001
0.3000	0.0000
0.5000	0.0000
0.7000	0.0000
0.9000	0.0000
1.0000	0.0000



Figur 2: median fitness for different mutation probabilites ($P_{mutation}$)

3.2.1 brief discussion

Roughly the optimal value of $P_{mutation}$ should be around 0.0175±0.015. Given that the GA has 50 genes per chromosome we see that $P_{mutation} \approx 0.02 = 1/50 = 1/(\#genes)$ However To decide clearly what $P_{mutation}$ should be both generally and for this case is hard to say given this data. To get an more accurate value, **(1)** the ammount of trials done to get an average should get increased, **(2)** smaller steps around $\mu = 0.0175$ should be taken, **(3)** Different ammount of genes should be tested (nGenes=25,50,100,200 are some good values), **(4)** The fitness function should be defined differently such that fitness $F(x^* + \delta x) \approx F(x^*)$. For this particular case $f(\boldsymbol{x}) = 1/g(\boldsymbol{x})$, thus any small change to $\boldsymbol{x}$ could make big difference to $F(\boldsymbol{x})$, thus making the data seem even more random.

3.3 C)

We make an educated guess given our data from A), $\rightarrow (x_1, x_2) = (3, 0.5)$. Now try to prove it analytically that it's a stationary point.
Applying chainrule when derivating:

$$\nabla g(x_1, x_2) = \mathbf{0} \rightarrow \begin{cases} 2(1.5 - x_1 + x_1 x_2^1)(x_2 - 1) + 2(2.25 - x_1 + x_1 x_2^2)(x_2^2 - 1) + 2(2.625 - x_1 + x_1 x_2^3)(x_2^3 - 1) = 0 \\ 2(1.5 - x_1 + x_1 x_2^1)(x_1) + 2(2.25 - x_1 + x_1 x_2^2)(2x_1 x_2) + 2(2.625 - x_1 + x_1 x_2^3)(3x_2^2 x_1) = 0 \end{cases}$$

$$\nabla g(x_1, x_2) = \begin{cases} P_1(\boldsymbol{x}) = (1.5 - x_1 + x_1 x_2^1) = \\ P_2(\boldsymbol{x}) = (2.25 - x_1 + x_1 x_2^2) \\ P_3(\boldsymbol{x}) = (2.625 - x_1 + x_1 x_2^3) \\ 2P_1(\boldsymbol{x})(x_2 - 1) + 2P_2(\boldsymbol{x})(x_2^2 - 1) + 2P_3(\boldsymbol{x})(x_2^3 - 1) = 0 \\ 2P_1(\boldsymbol{x})(x_1) + 2P_2(\boldsymbol{x})(2x_1 x_2) + 2P_3(\boldsymbol{x})(3x_2^2 x_1) = 0 \end{cases}$$

Evaluating $P_1(3,0.5), P_2(3,0.5), P_3(3,0.5)$

$$\begin{cases} P_1(3,0.5) = (1.5 - x_1 + x_1x_2) = 1.5 - 3 + 3/2 = 3 - 3 = 0 \\ P_2(3,0.5) = (2.25 - x_1 + x_1x_2^2) = 2.25 - 3 + 3/4 = 3 - 3 = 0 \\ P3(3,0.5) = (2.625 - x_1 + x_1x_2^3) = 2.625 - 3 + 3/8 = 3 - 3 = 0 \end{cases} \quad \rightarrow \nabla g(3,0.5) = \mathbf{0}$$

Thus proven it's a stationary point, as well as a global minimum since $g(x_1,x_2)$ only have squared terms and is of the form $P_1^2 + P_2^2 + P_3^3, P_1(3,0.5) = P_2(3,0.5) = P_3(3,0.5) = 0$