

Stochastic optimization algorithms 2018

Home problems, set 1

Cassard Sebastien 19960526-T313

September 25, 2018

1 Problem 1.1: Penalty method

1.1 Defining $f_p(X; \mu)$

Let's introduce the penalty function (penalty term) P defined by :

$$P(X; \mu) = \mu \left(\sum_{i=0}^m \max\{g_i(X), 0\}^2 + \sum_{i=0}^n h_i(X)^2 \right)$$

Where g_i , $i = 0 \dots m$ are inequality constraints and h_i , $i = 0 \dots n$ are equality constraints.

For the considered problem, with $X = (x_1, x_2)^T$:

$$P(x_1, x_2; \mu) = \mu (\max\{x_1^2 + x_2^2 - 1, 0\}^2)$$

Now we can define $f_p(x_1, x_2; \mu)$:

$$\begin{aligned} f_p(x_1, x_2; \mu) &= f(x_1, x_2) + P(x_1, x_2; \mu) \\ &= (x_1 - 1)^2 + 2(x_2 - 2)^2 + \mu (\max\{x_1^2 + x_2^2 - 1, 0\}^2) \end{aligned}$$

1.2 Computing $\nabla f_p(x_1, x_2; \mu)$

$$\nabla f_p(x_1, x_2; \mu) = \begin{pmatrix} \frac{\partial f_p}{\partial x_1} \\ \frac{\partial f_p}{\partial x_2} \end{pmatrix} = \begin{cases} \begin{pmatrix} 2(x_1 - 1) + 4\mu x_1(x_1^2 + x_2^2 - 1) \\ 4(x_2 - 2) + 4\mu x_2(x_1^2 + x_2^2 - 1) \end{pmatrix} & \text{if } (x_1^2 + x_2^2 - 1) > 0 \\ \begin{pmatrix} 2(x_1 - 1) \\ 4(x_2 - 2) \end{pmatrix} & \text{if } (x_1^2 + x_2^2 - 1) \leq 0 \end{cases}$$

1.3 Unconstrained minimum

To find the unconstrained minimum, we set $\mu = 0$. $\nabla f_p(x_1, x_2; \mu)$ is now given by :

$$\nabla f_p(x_1, x_2) = \begin{pmatrix} \frac{\partial f_p}{\partial x_1} \\ \frac{\partial f_p}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2(x_1 - 1) \\ 4(x_2 - 2) \end{pmatrix}$$

The unconstrained minimum can be found by solving $\nabla f_p(x_1, x_2; \mu) = (0, 0)^T$, which gives us the unconstrained minimum $X^* = (x_1, x_2)^T = (1, 2)^T$.

1.4 Program output

In the table below are given the results obtained by running the Penalty method algorithm with the following parameters:

- $\eta = 0.0001$
- $T = 10^{-6}$
- $\mu = [1, 10, 100, 1000]$

μ	x_1^*	x_2^*
1	0.434	1.210
10	0.331	0.996
100	0.314	0.955
1000	0.312	0.951

These results seem reasonable. The more μ increases, the more the values seem to converge towards the same point, approximately around $(0.31, 0.95)$.

2 Problem 1.2: Constrained Optimization

2.1 Analytical method

The set S is closed and limited, so it is compact. Thus we can use the analytical method described in the book p29-30 to solve this problem:

Determine the global minimum $(x_1^*, x_2^*)^T$ of f on the closed set S

$$f(x_1, x_2) = 4x_1^2 - x_1x_2 + 4x_2^2 - 6x_2$$

2.1.1 Finding stationary points in the interior of S

First let's find the stationary points in the interior of S by solving $\nabla f(x_1, x_2) = (0, 0)^T$:

$$\begin{aligned}\nabla f(x_1, x_2) &= \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 8x_1 - x_2 \\ -x_1 + 8x_2 - 6 \end{pmatrix} \\ \nabla f(x_1, x_2) = (0, 0)^T &\Leftrightarrow \begin{cases} 8x_1 - x_2 = 0 \\ -x_1 + 8x_2 - 6 = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} 8x_1 - x_2 = 0 \\ x_1 = 8x_2 - 6 \end{cases} \\ &\Leftrightarrow \begin{cases} 63x_2 - 48 = 0 \\ x_1 = 8x_2 - 6 \end{cases} \\ &\Leftrightarrow \begin{cases} x_2 = \frac{48}{63} = \frac{16}{21} \\ x_1 = 8\frac{16}{21} - 6 = \frac{2}{21} \end{cases}\end{aligned}$$

Let P_1 the stationary point of f found above :

$$P_1 = \left(\frac{2}{21}, \frac{16}{21} \right)^T$$

2.1.2 Exploring boundary

Now that we have explored the interior, let's examine the boundaries of the set S .

$$\text{(i)} : 0 < x_1 < 1, x_2 = 1 : f(x_1, 1) = 4x_1^2 - x_1 + 4 - 6 = 4x_1^2 - x_1 - 2$$

$$\frac{\partial f}{\partial x_1} = 8x_1 - 1$$

The stationary point is given by solving $\frac{\partial f}{\partial x_1} = 0$, and we find : $P_2 = (\frac{1}{8}, 1)^T$

$$\text{(ii)} : 0 < x_2 < 1, x_1 = 0 : f(0, x_2) = 4x_2^2 - 6x_2$$

$$\frac{\partial f}{\partial x_2} = 8x_2 - 6$$

The stationary point is given by solving $\frac{\partial f}{\partial x_2} = 0$, and we find : $P_3 = (0, \frac{3}{4})^T$

$$\text{(iii)} : x_1 = x_2 : f(x_1, x_1) = 7x_1^2 - 6x_1$$

$$\frac{\partial f}{\partial x_1} = 14x_1 - 6$$

The stationary point is given by solving $\frac{\partial f}{\partial x_1} = 0$, and we find : $P_4 = (\frac{3}{7}, \frac{3}{7})^T$

2.1.3 Exploring the corners

Now it only remains to check the values in the corners :

$$P_5 = (0, 0)^T, P_6 = (1, 1)^T, P_7 = (0, 1)^T$$

2.1.4 Investigating the points

The 7 values found above are the only ones that can be the overall minimum of f in Set S . Let's examine $P_i, i \in \{1...7\}$ successively:

- $f(P_1) \simeq -2.28$
- $f(P_2) \simeq -2.06$
- $f(P_3) \simeq -2.25$
- $f(P_4) \simeq -1.28$
- $f(P_5) \simeq 0$
- $f(P_6) \simeq 1$
- $f(P_7) \simeq -2$

From these value, $f(P_1)$ is the smallest. We can conclude that the global minimum $(x_1^*, x_2^*)^T$ of f on the closed set S is $(\frac{2}{21}, \frac{16}{21})^T$.

2.2 Lagrange Multiplier Method

Let's use the Lagrange Multiplier Method to determine the minimum $(x_1^*, x_2^*)^T$ of :

$$f(x_1, x_2) = 15 + 2x_1 + 3x_2$$

$$s.t : h(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2 - 21 = 0$$

First, we need to define $L(x_1, x_2, \lambda)$ as followed :

$$L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda h(x_1, x_2) = 15 + 2x_1 + 3x_2 + \lambda(x_1^2 + x_1x_2 + x_2^2 - 21)$$

We then have to find the stationary point of L , by solving :

$$\begin{cases} \frac{\partial L}{\partial x_1} = 0 \\ \frac{\partial L}{\partial x_2} = 0 \\ \frac{\partial L}{\partial \lambda} = 0 \end{cases}$$

Leading us to solve this system (As $\lambda = 0$ is not possible, we will now assume $\lambda \neq 0$):

$$\begin{cases} 2 + \lambda(2x_1 + x_2) = 0 \\ 3 + \lambda(2x_2 + x_1) = 0 \\ x_1^2 + x_1x_2 + x_2^2 - 21 = 0 \end{cases} \Leftrightarrow \begin{cases} x_2 = -\frac{2}{\lambda} - 2x_1 \\ x_1 = -\frac{3}{\lambda} - 2x_2 \\ x_1^2 + x_1x_2 + x_2^2 - 21 = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x_2 = -\frac{2}{\lambda} - 2x_1 \\ x_1 = -\frac{3}{\lambda} + 2(-\frac{2}{\lambda} - 2x_1) \\ x_1^2 + x_1x_2 + x_2^2 - 21 = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x_2 = -\frac{4}{3\lambda} \\ x_1 = -\frac{1}{3\lambda} \\ (-\frac{1}{3\lambda})^2 + \frac{1}{3\lambda} \frac{4}{3\lambda} + (-\frac{4}{3\lambda})^2 - 21 = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} x_2 = -\frac{4}{3\lambda} \\ x_1 = -\frac{1}{3\lambda} \\ \lambda \pm \frac{1}{3} \end{cases}$$

L seems to admit two stationary points $P_1 = (-1, -4, \frac{1}{3})^T$ and $P_2 = (1, 4, -\frac{1}{3})^T$. All that remains is to study these results to find the minimum of f : $f(-1, -4) = 1$ and $f(1, 4) = 29$. We can now say that the minimum $(x_1^*, x_2^*)^T$ of is $(-1, -4)^T$.

3 Basic GA program

3.1 Testing the GA

By running the *FunctionOptimization.m* script, the GA algorithm finds the minimum of function g at point $(0, -1)^{-T}$. This result was obtained with the following parameters :

- `populationSize = 30;`
- `numberOfGenes = 40;`
- `crossoverProbability = 0.8;`
- `mutationProbability = 0.025;`
- `tournamentSelectionParameter = 0.75;`
- `tournamentSize = 2;`
- `variableRange = 5.0;`
- `numberOfGenerations = 100;`

3.2 Analysis of the influence of the mutation rate

The data presented in the table below were measured using the *FunctionOptimizationMutationRateMeasure.m* script :

Mutation Rate Value	Median fitness value
0.0	0.0197
0.02	0.3230
0.05	0.3333
0.1	0.3333

We can notice that the more the mutation rate increases, the higher the fitness median is. The random mutation of genes is therefore an element that improves the search of the optimum. However, we can observe that the value of fitness stagnates very quickly from a mutation rate of 0.05. We can suppose that the mutation phenomenon is effective when it occurs at low frequencies. It can be assumed that for too high mutation rates, the median fitness is degraded. For maximum efficiency, very low mutation rates should be preferred.