A study of the spherical coordinates parameterization

Xiaoke Shen

the Graduate Center, the City University of New York

Abstract

This report provides the results of using the spherical coordinates to resolve the constrained problems.

1 Introduction

For the resource allocation problem, suppose resource p_i is allocated to i where i is the index of the object which get the resource, p_i is the allocation ratio to the total available resource. Then we can get:

$$\sum_{i=1}^{n} p_i = 1 \tag{1}$$

We can rewrite the equation 1 as bellow:

$$\sum_{i=1}^{n} (r_i)^2 = 1 \tag{2}$$

Any feasible allocation vector $r = (r_1, ..., r_n)$ on the unit ball can be described through $n-1$ angels denoted by $\theta_i, 1 \leq i \leq n-1$, in the following way. Indeed, the spherical coordinated parameterization of r via $\theta^T = (\theta_1, ..., \theta_{n-1})$ is given by:

$$r_i(\theta) = \begin{cases} cos(\theta_1), & i = 1; \\ cos(\theta_i)\prod_{k=1}^{i-1} sin(\theta_k), & 2 \leq i \leq n-1; \\ sin(\theta_{n-1})\prod_{k=1}^{n-2} sin(\theta_k), & i = n. \end{cases} \tag{3}$$

1.1 Penalty Methods

In order to adress the constrained optimization problem, several methods are provided. In this work, the penalty method is studied to compare the performance of the spherical coordinated parameterization method.
Given the constrained problem as described in equation 4

$$\begin{aligned} \min_{x\in\mathbb{R}^n} & J(x), \\ \text{s.t.} \quad & 0 \le x_i \le 1, i = 1, ..., n \\ & \sum_{i=1}^{n} x_i = 1 \end{aligned} \tag{4}$$

THe penalty methods modify the original performance function to penalize the extent to which the constraints are not satisfied. Recall that $|.|$ denotes the Euclidean norm, then the penalized function is defined:

$$\begin{aligned} J_\alpha(x) = J(x) + \frac{\alpha}{2}(|g(x)_+|^2 + |g(x)_-|^2), \\ where \quad g(x)_+ = (g_1(x)_+, ...g_j(x)_+)^T, and \quad g_i(x)_+ = max(1, g_i(\theta)) - 1 \\ where \quad g(x)_- = (g_1(x)_-, ...g_j(x)_-)^T, and \quad g_i(x)_- = min(0, g_i(\theta)) \end{aligned} \tag{5}$$

Then two-time scale algorithm is used to get the optimal value of x by using the formulas shown in equation 6, equation 7 and equation 8.

$$x_{n+1} = x_n - \epsilon_n \nabla J\alpha_n(x_n)^T \tag{6}$$

$$\alpha_{n+1} = \alpha_n + \delta_n \mathbb{1}_{\{n\in T_i\}} \tag{7}$$

where $\sum \delta_n = +\infty$, and

$$\nabla J\alpha_n(x_n) = \nabla_x J(x_n) + \alpha_n(g(x_n)^T \nabla g(x_n) \mathbb{1}_{\{|g(x_n)|<0\}} + g(x_n)^T \nabla g(x_n) \mathbb{1}_{\{|g(x_n)|>1\}}) \tag{8}$$

2 Deterministic Optimization

2.1 Cost function is linear function

2.1.1 Plots of the cost function

In this experiment, the cost function is a linear function. The optimization problem is described as below:

$$
\begin{aligned}
\min_{x_1, x_2 \in \mathbb{R}} \quad & J(x_1, x_2) = -2x_1 - x_2 \\
\text{s.t.} \quad & 0 \leq x_1 \leq 1 \\
& 0 \leq x_2 \leq 1 \\
& \sum_{i=1}^{2} x_i = 1
\end{aligned} \tag{9}
$$

As $\sum_{i=1}^{2} x_i = 1$, we can change equation 9 to:

$$
\begin{aligned}
\min_{x \in \mathbb{R}} \quad & J(x) = -x - 1 \\
\text{s.t.} \quad & 0 \leq x \leq 1
\end{aligned} \tag{10}
$$

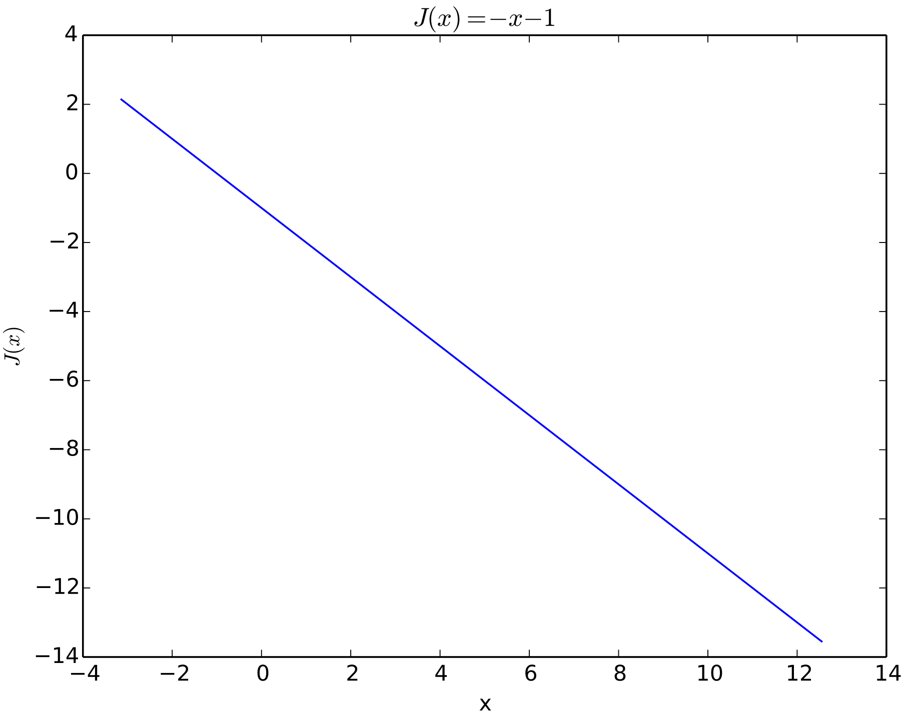

Figure 1: The plot of the cost function: $J(x) = -x - 1$.

$$
\min_{\theta \in \mathbb{R}} \quad J(\theta) = -cos(\theta)^2 - 1 \tag{11}
$$

The plot of the equation 10 is shown in figure 1. By using the equation 3, we can change the constrained optimization problem to an unconstrained problem by using the spherical coordinates parameterization method. The unconstrained problem by using the spherical coordinates parameterization method is formally described in equation 11. The plot of the equation 11 is shown in figure 2.

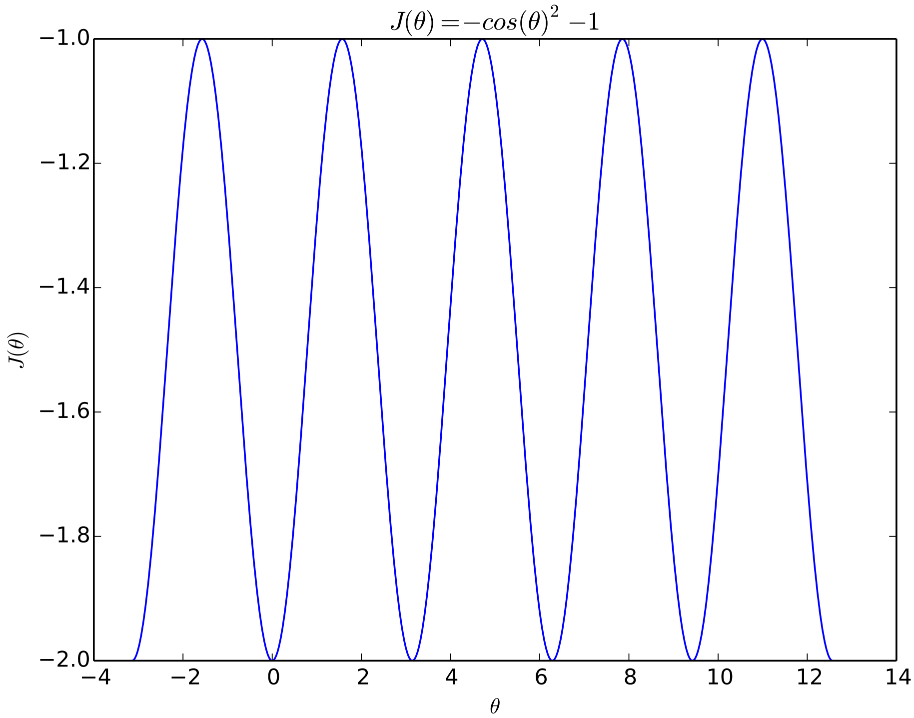

Figure 2: The plot of the cost function: $J(\theta) = -cos(\theta)^2 - 1$.

2.1.2 Penalty method

By using the equation 6, equation 7 and equation 8, for this specified question, we can get:

$$x_{n+1} = x_n - \epsilon_n \nabla J \alpha_n (x_n)^T \tag{12}$$

$$\alpha_{n+1} = \alpha_n + \delta_n \mathbb{1}_{\{n \in T_i\}} \tag{13}$$

$$\nabla_x J(x_n) = -1 \tag{14}$$

$$\nabla J\alpha_n(x_n) = \nabla_x J(x_n) + \alpha_n x_n(\mathbb{1}_{\{x_n<0\}} + x_n\mathbb{1}_{\{x_n>1\}}) \tag{15}$$

By using the equation 12 , equation 13, equation 14 and equation 15, we can estimate the optimal value of x and the estimation process is shown in figure 3

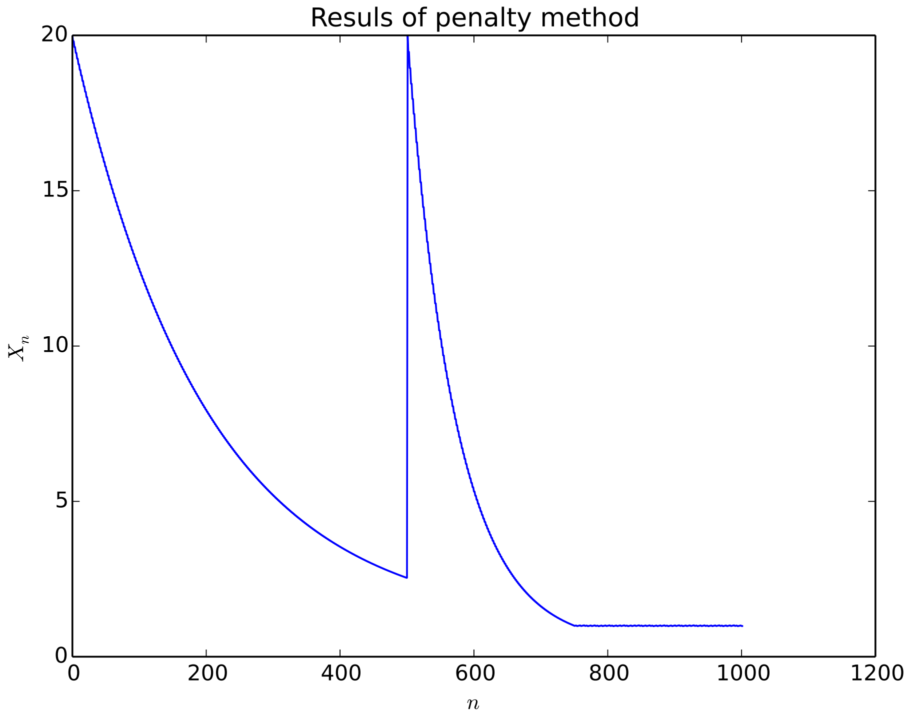

Figure 3: The estimation process of the optimal x when the penalty method is used.Here the initial value of x is set to 20. The total iteration for T_i is 500. The $\alpha_1 = 1$ and $\alpha_{n+1} = \alpha_n + \delta_n$, where $\delta_n = e^n$.

2.1.3 the spherical coordinated parameterization method

$$\theta_{n+1} = \theta_n - \epsilon_n \nabla J\theta_n \tag{16}$$

$$\nabla_\theta J(\theta_n) = 2 * cos(\theta_n) * sin(\theta_n) \tag{17}$$

When the spherical coordinated parameterization is used, the constrained optimization problem can be changed into a unconstrained problem. We can use the formula shown in equation 16 and equation 17 to estimate the optimal result. As we don't know whether the cost function is convex, we may get multiple local minimum. In order to get a global minimum, several experiments are done to get more than one estimation and the best one will be selected as the global optimization. In this experiment, 5 experiments are done. The estimation process is shown in figure 4.

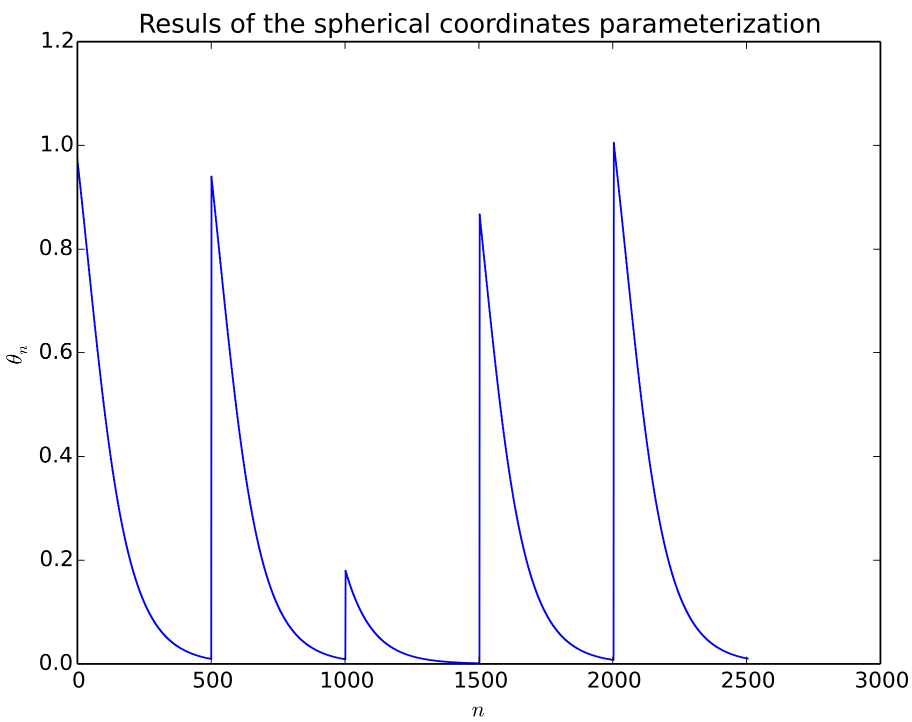

Figure 4: The estimation process of the optimal θ when the spherical coordinated parameterization is used. In order to make sure the global optimal value is collected, 5 times experiments with random initialized value of θ is done and the global minimum value is collected from those 5 experiments.

2.1.4 Optimization results comparison

The comparison of the penalty and the spherical coordinates parameterization method is shown in table 1. From the results, we can see that the spherical coordinates parameterization method can get a more accurate result. At the same time, the spherical coordinates parameterization method

Optimal x	method	estimated x	execution time(ms)
1.0	penalty method	0.995	0.71
1.0	the spherical coordinates parameterization	1.000	9.55

Table 1: The results of the penalty method and spherical coordinate parameterization.

is slower as 5 repeated experiments are done to get the global minimum result.

2.2 Cost function is polynomial

2.2.1 Plots of the cost function

In this experiment, the cost function is a linear function. The optimization problem is described as below:

$$\begin{aligned} \min_{x\in\mathbb{R}} \quad & J(x) = -x^4 + 2x^2 - x \\ \text{s.t.} \quad & 0 \leq x \leq 1 \end{aligned} \tag{18}$$

By using this cost function, we can get a minimum value which is between 0 and 1. We can change the problem as below by using the spherical coordinates parameterization method.

$$\min_{\theta\in\mathbb{R}} \quad J(\theta) = -cos(\theta)^8 + 2cos(\theta)^4 - cos(\theta)^2 \tag{19}$$

The plot of the equation 18 is shown in figure 5. The plot of the equation 11 is shown in figure 6.

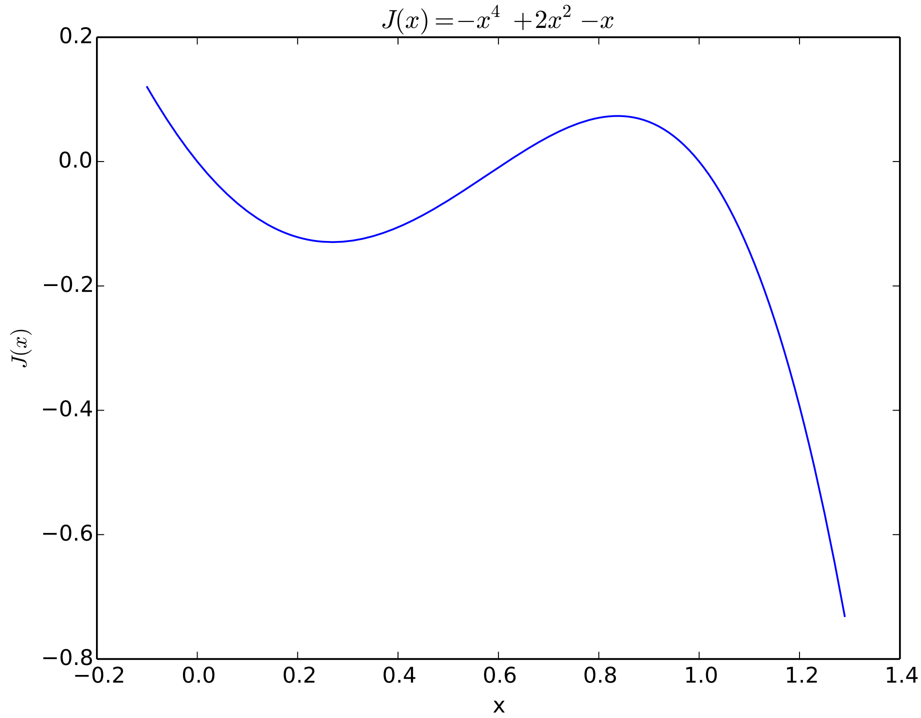

Figure 5: The plot of the cost function: $J(x) = -x^4 + 2x^2 - x$.

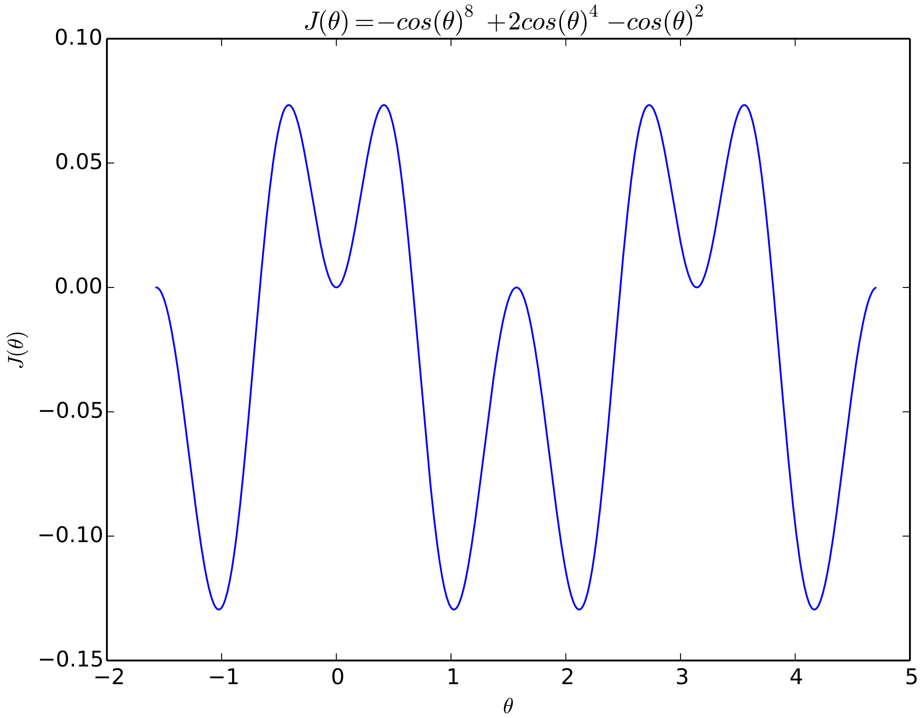

Figure 6: The plot of the cost function: $J(\theta) = -cos(\theta)^8 + 2cos(\theta)^4 - cos(\theta)^2$.

2.2.2 Penalty method

By using the equation 6, equation 7 and equation 8, for this specified question, we can get:

$$x_{n+1} = x_n - \epsilon_n \nabla J\alpha_n(x_n)^T \tag{20}$$

$$\alpha_{n+1} = \alpha_n + \delta_n \mathbb{1}_{\{n \in T_i\}} \tag{21}$$

$$\nabla_x J(x_n) = -4x^3 + 4x - 1 \tag{22}$$

$$\nabla J\alpha_n(x_n) = \nabla_x J(x_n) + \alpha_n x_n (\mathbb{1}_{\{x_n < 0\}} + x_n \mathbb{1}_{\{x_n > 1\}}) \tag{23}$$

By using the equation 20 , equation 21, equation 22 and equation 23, we can estimate the optimal value of x and the estimation process is shown in figure 7

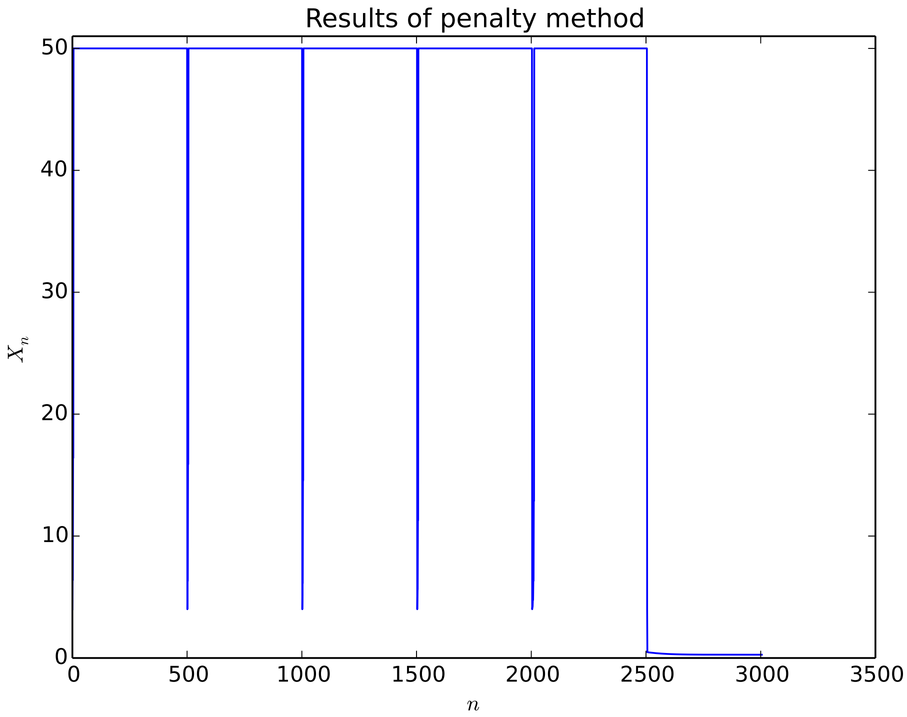

Figure 7: The estimation process of the optimal x when the penalty method is used.Here the initial value of x is set to 4. The total iteration for T_i is 500. The $\alpha_1 = 1$ and $\alpha_{n+1} = \alpha_n + \delta_n$, where $\delta_n = e^n$ The value of x_n is upper bounded by the value of 50 as if it is not bounded, the value of x_n will become too large to be handled by the computer.

2.2.3 the spherical coordinated parameterization method

$$\theta_{n+1} = \theta_n - \epsilon_n \nabla J \theta_n \tag{24}$$

$$\nabla_\theta J(\theta_n) = 8 * cos(\theta_n)^7 * sin(\theta_n) - 8 * cos(\theta_n)^3 * sin(\theta_n) + 2 * cos(\theta_n) * sin(\theta_n) \tag{25}$$

When the spherical coordinated parameterization is used, the constrained optimization problem can be changed into a unconstrained problem. We can use the formula shown in equation **??** and equation **??** to estimate the optimal result. As we don't know whether the cost function is convex, we may get multiple local minimum. In order to get a global minimum, several experiments are done to get more than one estimation and the best one will be selected as the global optimization. In this experiment, 5 experiments are done. The estimation process is shown in figure ??.

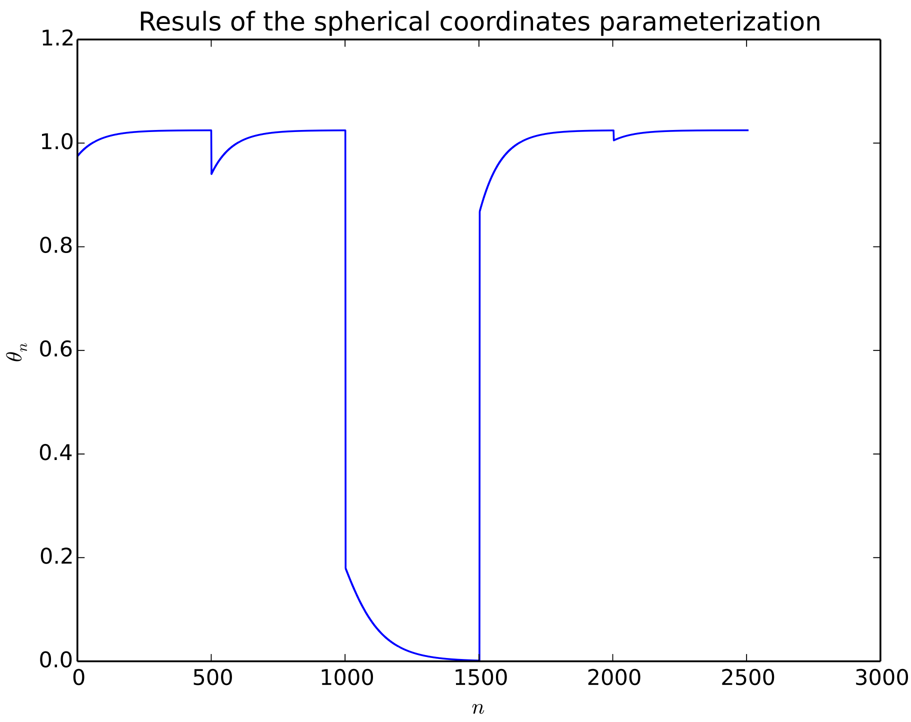

Figure 8: The estimation process of the optimal θ when the spherical coordinated parameterization is used. In order to make sure the global optimal value is collected, 5 times experiments with random initialized value of θ is done and the global minimum value is collected from those 5 experiments.

2.2.4 Optimization results comparison

Optimal x	method	estimated x	execution time(ms)
0.2696	penalty method	0.2697	2.34
0.2696	the spherical coordinates parameterization	0.2696	12.11

Table 2: The results of the penalty method and spherical coordinate parameterization.

The comparison of the penalty and the spherical coordinates parameterization method is shown in table ??. We can get the similar conclusion as the cost function is linear.

2.3 Cost function is a combination of the log normal pdf function and the exponential function

2.3.1 Plots of the cost function

In this experiment, the cost function is a combination of the log normal pdf function and the exponential function. The optimization problem is described as below:

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & J(x) = -\frac{1}{(x+0.001)\sqrt{2\pi}} e^{-\frac{(\ln(x+0.001))^2}{2}} - 0.01e^{x} \\ \text{s.t.} \quad & 0 \leq x \leq 1 \end{aligned} \tag{26}$$

By using this cost function, we can get a minimum value which is between 0 and 1. We can change the problem as below by using the spherical coordinates parameterization method.

$$\min_{\theta \in \mathbb{R}} \quad J(\theta) = -\frac{1}{(\cos\theta^2 + 0.001)\sqrt{2\pi}} e^{-\frac{\ln(\cos\theta^2+0.001)^2}{2}} - 0.01e^{\cos\theta^2} \tag{27}$$

The plot of the equation ?? is shown in figure ??. The plot of the equation ?? is shown in figure 6.

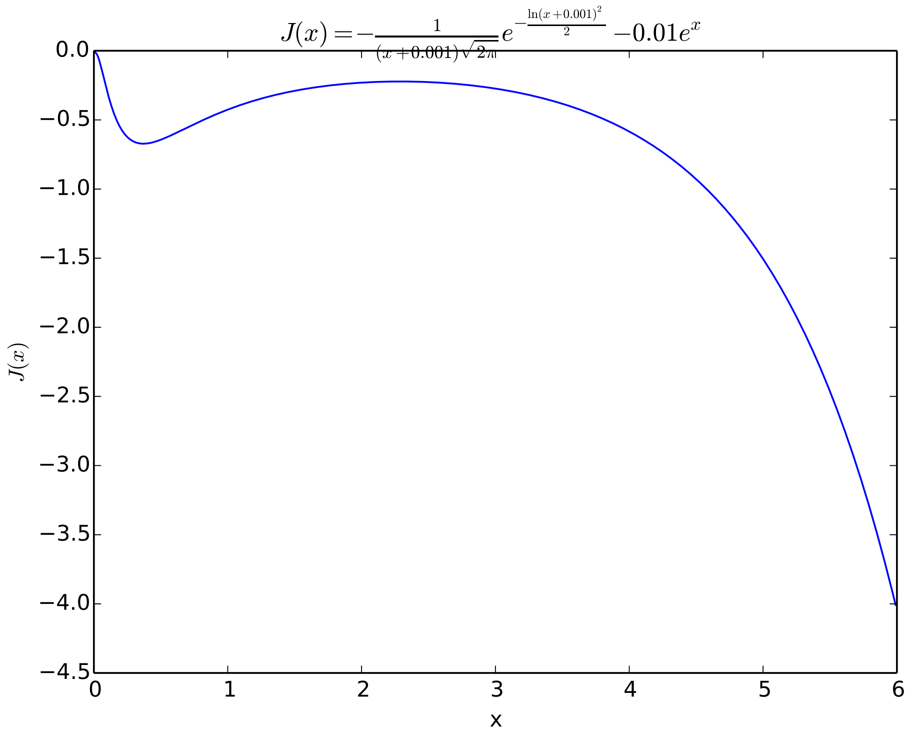

Figure 9: The plot of the cost function: $J(x) = \frac{1}{(x+0.001)\sqrt{2\pi}} e^{-\frac{\ln(x+0.001)^2}{2}} - 0.01e^x$.

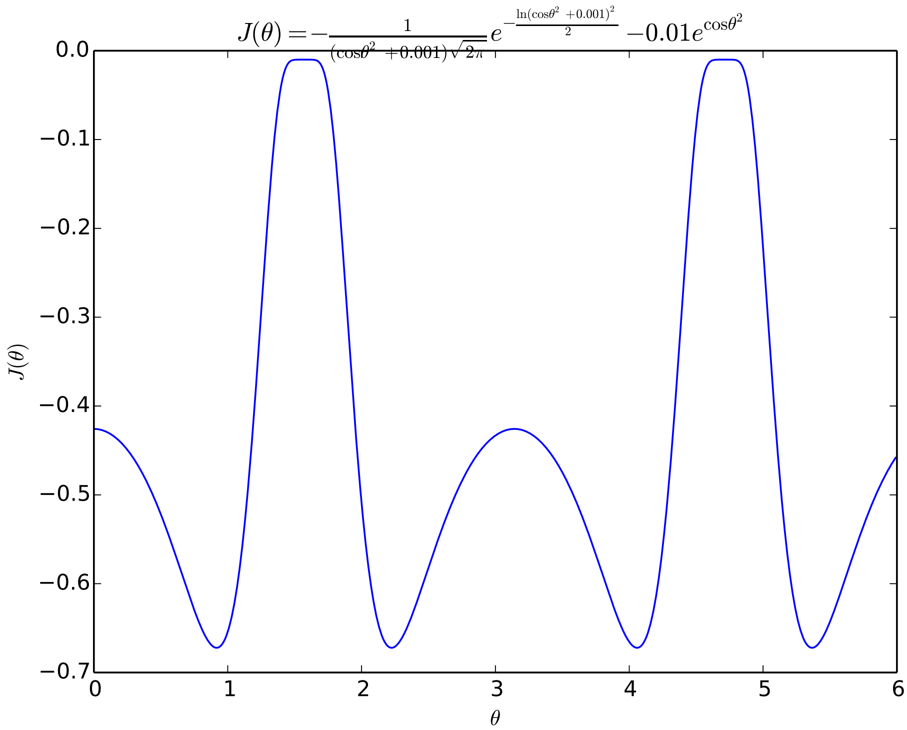

Figure 10: The plot of the cost function: $J(\theta) = -\frac{1}{(\cos\theta^2+0.001)\sqrt{2\pi}} e^{-\frac{\ln(\cos\theta^2+0.001)^2}{2}} - 0.01e^{\cos\theta^2}$.

2.3.2 Penalty method

By using the equation 6, equation 7 and equation 8, for this specified question, we can get:

$$x_{n+1} = x_n - \epsilon_n \nabla J\alpha_n(x_n)^T \tag{28}$$

$$\alpha_{n+1} = \alpha_n + \delta_n \mathbb{1}_{\{n \in T_i\}} \tag{29}$$

$$\nabla_x J(x) = -\frac{1 + \ln(x + 0.001)}{(x + 0.001)^2\sqrt{2\pi}} e^{-\frac{\ln(x+0.001)^2}{2}} - 0.01e^x \tag{30}$$

$$\nabla J\alpha_n(x_n) = \nabla_x J(x_n) + \alpha_n x_n(\mathbb{1}_{\{x_n<0\}} + x_n\mathbb{1}_{\{x_n>1\}}) \tag{31}$$

By using the equation **??** , equation **??**, equation **??** and equation **??**, we can estimate the optimal value of x and the estimation process is shown in figure 7

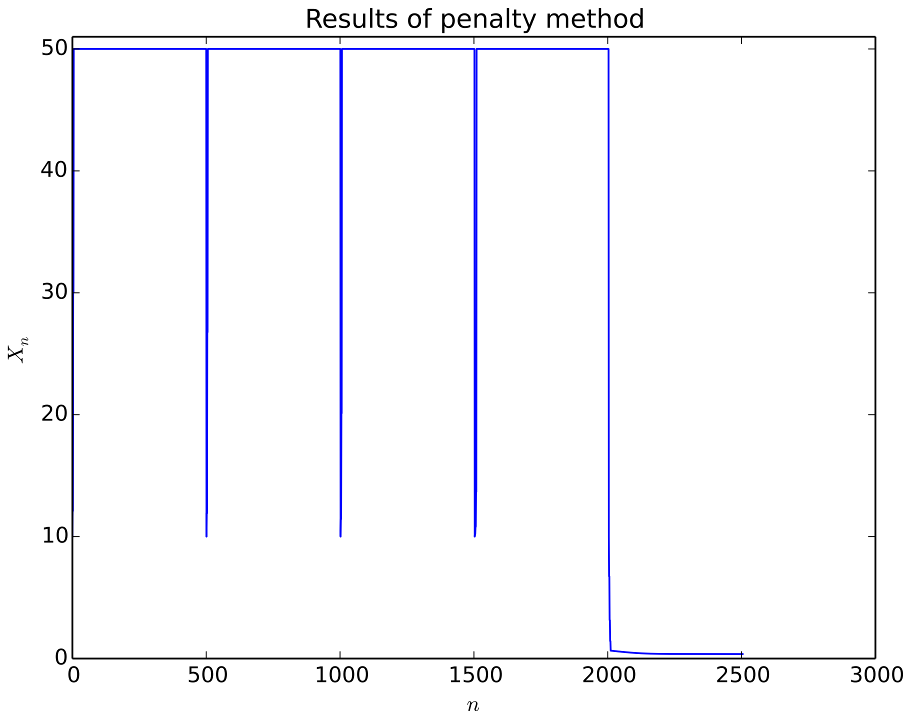

Figure 11: The estimation process of the optimal x when the penalty method is used.Here the initial value of x is set to 10. The total iteration for T_i is 500. The $\alpha_1 = 1$ and $\alpha_{n+1} = \alpha_n + \delta_n$, where $\delta_n = e^n$ The value of x_n is upper bounded by the value of 50 as if it is not bounded, the value of x_n will become too large to be handled by the computer.

2.3.3 the spherical coordinated parameterization method

$$\theta_{n+1} = \theta_n - \epsilon_n \nabla J \theta_n \tag{32}$$

$$\nabla_\theta J(\theta_n) = \left(\frac{1 + \ln(\cos\theta^2 + 0.001)}{(\cos\theta^2 + 0.001)^2\sqrt{2\pi}} e^{-\frac{\ln(\cos\theta^2+0.001)^2}{2}} - 0.01 e^{\cos\theta^2}\right)(-2\cos\theta\sin\theta) \tag{33}$$

When the spherical coordinated parameterization is used, the constrained optimization problem can be changed into a unconstrained problem. We can use the formula shown in equation **??** and equation **??** to estimate the optimal result. As we don't know whether the cost function is convex, we may get multiple local minimum. In order to get a global minimum, several experiments are done to get more than one estimation and the best one will be selected as the global optimization. In this experiment, 5 experiments are done. The estimation process is shown in figure ??.

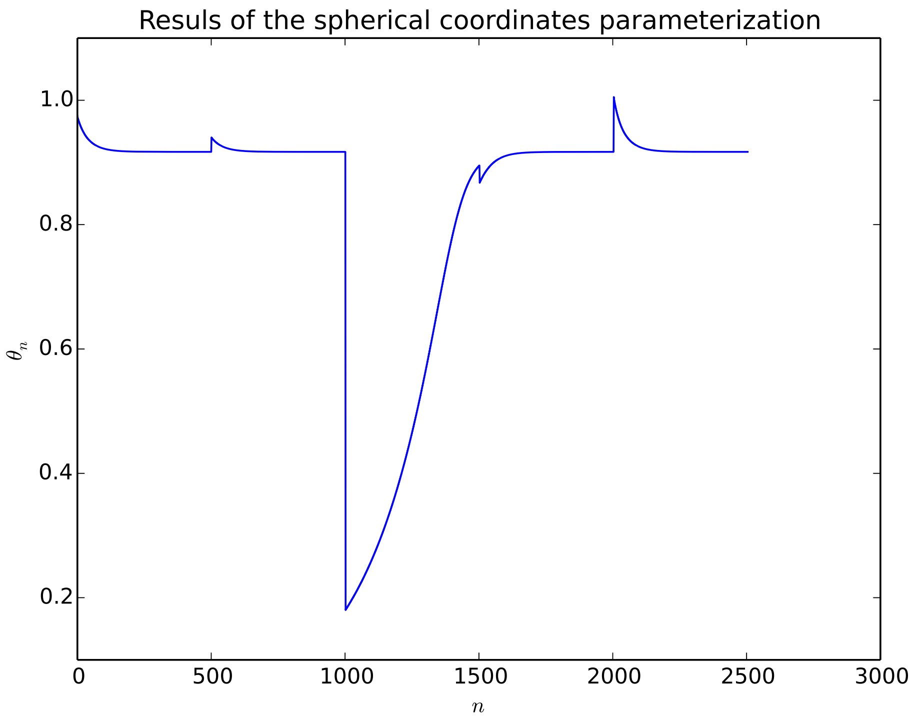

Figure 12: The estimation process of the optimal θ when the spherical coordinated parameterization is used. In order to make sure the global optimal value is collected, 5 times experiments with random initialized value of θ is done and the global minimum value is collected from those 5 experiments.

2.3.4 Optimization results comparison

The comparison of the penalty and the spherical coordinates parameterization method is shown in table ??.

Optimal x	method	estimated x	execution time(ms)
0.3671	penalty method	0.3699	4.64
0.3671	the spherical coordinates parameterization	0.3699	24.28

Table 3: The results of the penalty method and spherical coordinate parameterization.

3 Undeterministic Optimization by adding the gaussian noise to the gradient

3.1 Cost function is polynomial

The cost function is the same as equation 18. The noise is added to the gradient of the cost function as below:

$$\begin{aligned} \nabla_x J(x_n) &= -4x^3 + 4x - 1 + \xi \\ &\textit{where} \quad \xi \sim \mathcal{N}(0,\, 1)\,. \end{aligned} \tag{34}$$