Linear and Convex Optimization Homework 11

Qiu Yihang, 2021/12/12-12/13

0. Preparation2

Complete proj_gd.py. The completed code (with `proj_gd` function) is enclosed in the zip file.

1. *Solution:*

I made some adjustments to the given p1.py and utils.py. The code is enclosed in the zip file.

Set step size as 0.1.

Use Projected Gradient Descent with initial point $\boldsymbol{x}_0 = (-1,0.5)^T$.

The solution and the number of iterations is given below.

```
In [2]: runfile('D:/Textbooks/2021-2022-1/Linear and Convex
Optimization/hw11/P1.py', wdir='D:/Textbooks/2021-2022-1/Linear and
Convex Optimization/hw11')

t =  1
stepsize =  0.1
number of iterations: 69
solution: [9.99999997e-01 2.52227830e-09]
value: 4.5
```

Fig.01. Results of Program 1

Ignoring numerical errors, the solution of projected gradient descent should be

$$\boldsymbol{x}^\star = (1,0)^T, f(\boldsymbol{x}^\star) = \frac{9}{2}.$$

The visualization of the trajectory of x_k and the change of error $f(\boldsymbol{x}_k) - f(\boldsymbol{x}^\star)$ are as follows.

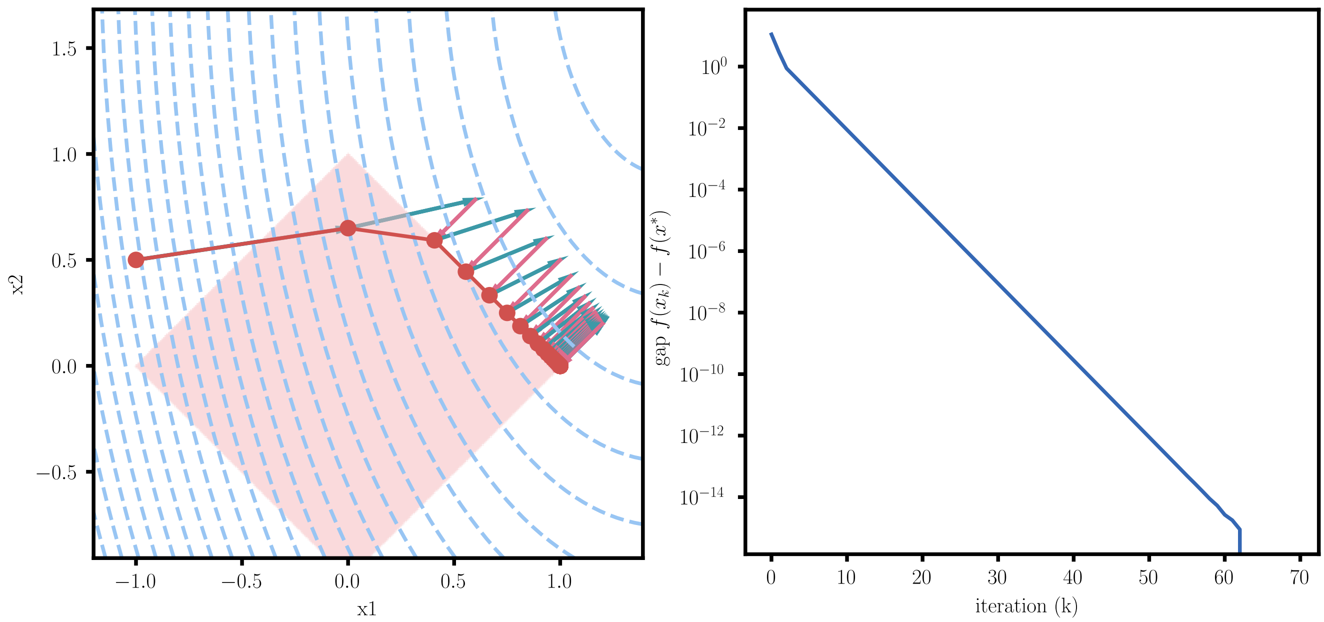

Fig.02. Trajectory of $\boldsymbol{x}_k$ Produced by Projected Gradient Descent

and Change of Gap $f(\boldsymbol{x}_k) - f(\boldsymbol{x}^\star)$

2.(a) ***Solution:***

The Lagrangian function is

$$\mathcal{L}(\boldsymbol{x},\lambda) = e^{x_1} + e^{2x_2} + e^{2x_3} + \lambda(x_1 + x_2 + x_3 - 1).$$

Let the optimal solution be $\boldsymbol{x}^\star$ with corresponding Lagrangian multiplier $\lambda^\star$. Thus, we have

$$\begin{cases} \nabla\mathcal{L}_{x_1}(\boldsymbol{x}^\star,\lambda^\star) = e^{x_1^\star} + \lambda^\star = 0 \\ \nabla\mathcal{L}_{x_2}(\boldsymbol{x}^\star,\lambda^\star) = 2e^{2x_2^\star} + \lambda^\star = 0 \\ \nabla\mathcal{L}_{x_3}(\boldsymbol{x}^\star,\lambda^\star) = 2e^{2x_3^\star} + \lambda^\star = 0 \\ \nabla\mathcal{L}_{\lambda}\ (\boldsymbol{x}^\star,\lambda^\star) = x_1^\star + x_2^\star + x_3^\star - 1 = 0 \end{cases} \Rightarrow \begin{cases} \boldsymbol{x}^\star = \left(\dfrac{1+\ln 2}{2}, \dfrac{1-\ln 2}{4}, \dfrac{1-\ln 2}{4}\right) & \blacksquare \\ \lambda^\star = -\sqrt{2\mathrm{e}} & \blacksquare \end{cases}$$

The optimal value is $f^\star = f(\boldsymbol{x}^\star) = 2\sqrt{2\mathrm{e}}$. $\blacksquare$

(b) ***Proof:***

First we calculate the projection onto an affine space $X = \{\boldsymbol{x}: \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$.

$$\mathcal{P}_X(\boldsymbol{x}) = \underset{\boldsymbol{y}\in X}{\operatorname{argmin}} \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 \qquad (\star)$$

The Lagrangian of Problem $(\star)$ is

$$\mathcal{L}(\boldsymbol{y},\boldsymbol{\lambda}) = \|\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \boldsymbol{\lambda}(\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}) = (\boldsymbol{x} - \boldsymbol{y})^T(\boldsymbol{x} - \boldsymbol{y}) + \boldsymbol{\lambda}(\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b})$$

Let the optimal solution be $\boldsymbol{y}^\star$ with corresponding Lagrangian multiplier $\boldsymbol{\lambda}^\star$. Thus, we have

$$\begin{cases} \nabla\mathcal{L}_{\boldsymbol{y}}(\boldsymbol{y}^\star,\boldsymbol{\lambda}^\star) = 2\boldsymbol{y}^\star - \boldsymbol{2x} + \boldsymbol{A}^T\boldsymbol{\lambda}^{\star T} = \boldsymbol{0} \\ \nabla\mathcal{L}_{\lambda}(\boldsymbol{y}^\star,\boldsymbol{\lambda}^\star) = \boldsymbol{A}\boldsymbol{y}^\star - b = 0 \end{cases} \Rightarrow \begin{cases} \boldsymbol{\lambda}^\star = -2\left((\boldsymbol{A}\boldsymbol{A}^T)^{-1}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})\right)^T \\ \boldsymbol{y}^\star = \boldsymbol{x} + \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}) \end{cases}$$

Thus,

$$\mathcal{P}_X(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}).$$

Based on the analysis above, I made some adjustments to the given `p2.py` (enclosed in the zip file).

Use Projected Gradient Descent to solve the problem numerically with initial point $\boldsymbol{x}_0 = (0,0,0)$. Set step size as 0.1. The solution and the number of iterations is given below. $\blacksquare$

```
In [1]: runfile('D:/Textbooks/2021-2022-1/Linear and Convex
Optimization/hw11/P2.py', wdir='D:/Textbooks/2021-2022-1/Linear and
Convex Optimization/hw11')

number of iterations: 52
solution: [[0.84657359 0.07671321 0.07671321]]
value: 4.663287963194248
```

Fig.03. Results of Program 2

Ignoring numerical errors and given that

$$\frac{1+\ln 2}{2} \approx 0.84657359, \frac{1-\ln 2}{4} \approx 0.07671320, 2\sqrt{2\mathrm{e}} \approx 4.663287963,$$

the result of the program is very approximate to the optimal solution calculated by hand.