Homework 6

Zhen

November 5, 2021

Problem 1

(a). Considering that $\nabla^2 f = Q$, the eigenvalues of $\nabla^2 f$ is γ and 1.
So we have $\max\{\gamma, 1\} \le |L|$. Then the smallest L such that f is L-smooth is $\max\{\gamma, 1\}$.

(b). $f(\boldsymbol{x}) - \frac{m}{2}\|\boldsymbol{x}\|^2 = \frac{\gamma - m}{2}x_1^2 + \frac{1-m}{2}x_2^2$ is convex. Thus $diag(\frac{\gamma - m}{2}, \frac{1-m}{2}) \succeq \boldsymbol{O}$.
Finally we have $m \le \min(\gamma, 1)$.

(c). When the step size is 1, 0.1 and 0.01, $\boldsymbol{x}_0$ converges. While for the size 2.2, it does not converge.

step sizes	Num of Iter	step sizes	Num of Iter
2.2	NaN	1	88
0.1	917	0.01	9206

Table 1: Number of Iteration

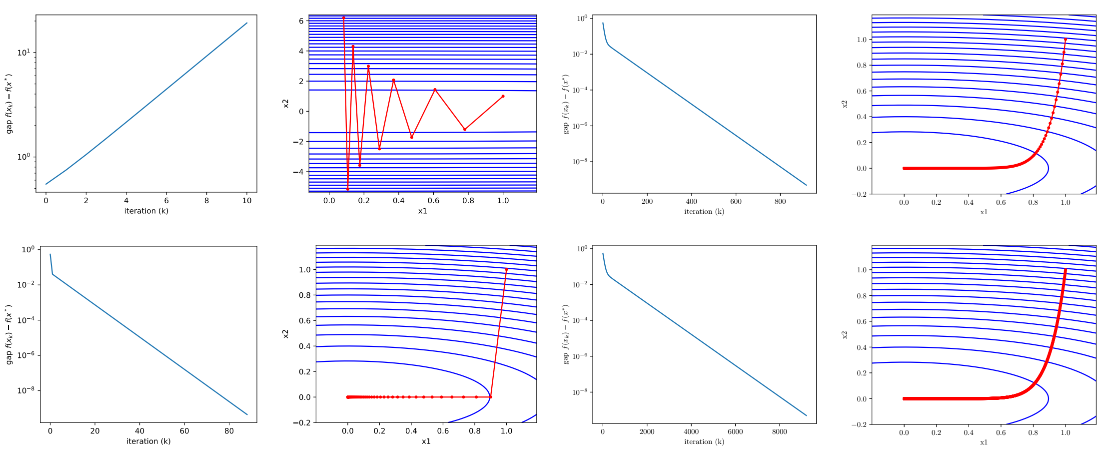


Figure 1: The step size of each case are:
upper left: 2.2, upper right 0.1, bottom left: 1.0, bottom right: 0.01

The output is:

```
gamma=0.1, stepsize=[1, 0.1, 0.01], number of iterations=88
gamma=0.1, stepsize=[1, 0.1, 0.01], number of iterations=917
gamma=0.1, stepsize=[1, 0.1, 0.01], number of iterations=9206
```

Figure 2: The Output of P1(c)

(d). The output of each case is as follows.

```
gamma=1, stepsize=1, number of iterations=1
gamma=0.1, stepsize=1, number of iterations=88
gamma=0.01, stepsize=1, number of iterations=688
gamma=0.001, stepsize=1, number of iterations=4603
```

By the figure we have: when $\gamma \downarrow$, iterations $\uparrow$.

Problem 2

Let the stepsize be 0.2 and finally we get:

```
GD: stepsize=0.2, number of iterations=26, solution = 1.4999991470913592, 1.9999982941827184
np.linalg.solve: solution = 1.5, 2.0
```

Figure 3: The Output of P2

Answer: $\boldsymbol{\omega} = (1.5000, 2.0000)^{\mathcal{T}}$.

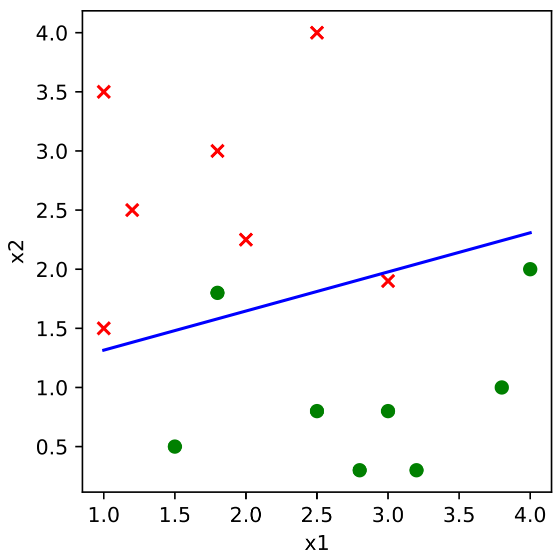
Apparently they agree with each other.

Problem 3

The output is:

```
accuracy = 0.8666666666666667
solution is [-1.47020052  4.44377575 -4.37548225]
```

Then we could get the visualization



Answer: $\boldsymbol{\omega} = (-1.470, 4.444, -4.375)^{\mathcal{T}}$, $Acc = 0.8667$

Problem 4

Suppose $f(\boldsymbol{x})$ is differentiable and α-strongly convex, and $g(\boldsymbol{x})$ is β-smooth. Show that the function $h(\boldsymbol{x}) = f(\boldsymbol{x}) - g(\boldsymbol{x})$ is convex if $\alpha \geq \beta$.

Proof:

By the premise we have:

$f(\boldsymbol{x}) - \frac{\alpha}{2}\|\boldsymbol{x}\|^2$ is convex which means that

$$f(\boldsymbol{x}) - \frac{\alpha}{2}\|\boldsymbol{x}\|^2 \geq f(\boldsymbol{y}) - \frac{\alpha}{2}\|\boldsymbol{y}\|^2 + (\nabla f(\boldsymbol{y}) - \alpha\boldsymbol{y})^{\mathcal{T}}(\boldsymbol{x} - \boldsymbol{y})$$

Also, because h is β-smooth,

$$h(\boldsymbol{x}) \leq h(\boldsymbol{y}) + \nabla h(\boldsymbol{y})^{\mathcal{T}}(\boldsymbol{x} - \boldsymbol{y}) + \frac{\beta}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2$$

Rearrange 2 above equations:

$$\nabla f(\boldsymbol{y})^{\mathcal{T}}(\boldsymbol{x} - \boldsymbol{y}) \leq f(\boldsymbol{x}) - f(\boldsymbol{y}) - \frac{\alpha}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2 \tag{1}$$

and

$$-\nabla h(\boldsymbol{y})^{\mathcal{T}}(\boldsymbol{x} - \boldsymbol{y}) \leq -h(\boldsymbol{x}) + h(\boldsymbol{y}) + \frac{\beta}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2 \tag{2}$$

Plugging Eq(1) and Eq(2):

$$\begin{aligned}(\nabla f(\boldsymbol{y}) - \nabla g(\boldsymbol{y}))^{\mathcal{T}}(\boldsymbol{x} - \boldsymbol{y}) &\leq f(\boldsymbol{x}) - g(\boldsymbol{x}) - f(\boldsymbol{y}) + g(\boldsymbol{y}) + \frac{\beta - \alpha}{2}\|\boldsymbol{x} - \boldsymbol{y}\|^2 \\ &\leq f(\boldsymbol{x}) - g(\boldsymbol{x}) - f(\boldsymbol{y}) + g(\boldsymbol{y})\end{aligned}$$

Rearrange the above Eq.:

$$\begin{aligned}& f(\boldsymbol{x}) - g(\boldsymbol{x}) \geq f(\boldsymbol{y}) - g(\boldsymbol{y}) + (\nabla f(\boldsymbol{y}) - \nabla g(\boldsymbol{y}))^{\mathcal{T}}(\boldsymbol{x} - \boldsymbol{y}) \\ \iff & h(\boldsymbol{x}) \geq h(\boldsymbol{y}) + \nabla h(\boldsymbol{y})^{\mathcal{T}}(\boldsymbol{x} - \boldsymbol{y}) \\ \iff & h(\boldsymbol{x}) \text{ is convex.}\end{aligned}$$