Optimization

Fall semester 2022/23

Kristóf Bérczi,
Lydia Mendoza

Eötvös Loránd University
Institute of Mathematics
Department of Operations Research

Outline

Question 1

Let $G = (V, E)$ be an undirected graph and $s, t \in V$. Consider the following problem:

$$\min \sum_{uv \in E} |x_u - x_v| \tag{1}$$

$$\text{s.t. } x_s - x_t = 1 \tag{2}$$

This is not a linear program in this form. Rewrite it as a linear program. (1pt)

Question 1

Let $G = (V, E)$ be an undirected graph and $s, t \in V$. Consider the following problem:

$$\min \sum_{uv \in E} |x_u - x_v| \tag{1}$$

$$\text{s.t. } x_s - x_t = 1 \tag{2}$$

This is not a linear program in this form. Rewrite it as a linear program. (1pt)

For each $uv \in E$ we can write $|x_u - x_v| = y_{uv}$ as: $\begin{cases} x_u - x_v \leqslant y_{uv} \\ -x_u + x_v \leqslant y_{uv} \end{cases}$

Question 1

Let $G = (V, E)$ be an undirected graph and $s, t \in V$. Consider the following problem:

$$\min \sum_{uv \in E} |x_u - x_v| \tag{1}$$

$$\text{s.t. } x_s - x_t = 1 \tag{2}$$

This is not a linear program in this form. Rewrite it as a linear program. (1pt)

For each $uv \in E$ we can write $|x_u - x_v| = y_{uv}$ as: $\begin{cases} x_u - x_v \leqslant y_{uv} \\ -x_u + x_v \leqslant y_{uv} \end{cases}$

Then we have the LP:

$$\min \sum_{uv \in E} y_{uv} \tag{3}$$

$$\text{s.t. } x_s - x_t = 1 \tag{4}$$

$$x_u - x_v \leqslant y_{uv} \quad \forall uv \in E \tag{5}$$

$$-x_u + x_v \leqslant y_{uv} \quad \forall uv \in E \tag{6}$$

$$y_{uv} \geqslant 0 \quad \forall uv \in E \tag{7}$$

Question 2, f_1

$f_1(w_1, w_2) = \frac{1}{2}w_1^2 + \frac{7}{2}w_2^2$

a. Calculate the gradients of the functions. (2pts)
b. Are these function convex? (2pts)
c. Determine the global minimum of the functions. (2pts)
d. Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

Question 2, f_1

$f_1(w_1, w_2) = \frac{1}{2}w_1^2 + \frac{7}{2}w_2^2$

- (a) Calculate the gradients of the functions. (2pts)
- (b) Are these function convex? (2pts)
- (c) Determine the global minimum of the functions. (2pts)
- (d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

(a) $\nabla f_1 = (\frac{\partial f_1}{\partial w_1}, \frac{\partial f_1}{\partial w_2}) = (w_1, 7w_2)$

Question 2, f_1

$f_1(w_1, w_2) = \frac{1}{2}w_1^2 + \frac{7}{2}w_2^2$

- (a) Calculate the gradients of the functions. (2pts)
- (b) Are these function convex? (2pts)
- (c) Determine the global minimum of the functions. (2pts)
- (d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

(a) $\nabla f_1 = (\frac{\partial f_1}{\partial w_1}, \frac{\partial f_1}{\partial w_2}) = (w_1, 7w_2)$

(b) $\nabla^2 f_1 = \begin{bmatrix} \frac{\partial^2 f_1}{\partial w_1^2} & \frac{\partial^2 f_1}{\partial w_1 w_2} \\ \frac{\partial^2 f_1}{\partial w_2 w_1} & \frac{\partial^2 f_1}{\partial w_2^2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$

Question 2, f_1

$f_1(w_1, w_2) = \frac{1}{2}w_1^2 + \frac{7}{2}w_2^2$

- (a) Calculate the gradients of the functions. (2pts)
- (b) Are these function convex? (2pts)
- (c) Determine the global minimum of the functions. (2pts)
- (d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

(a) $\nabla f_1 = (\frac{\partial f_1}{\partial w_1}, \frac{\partial f_1}{\partial w_2}) = (w_1, 7w_2)$

(b) $\nabla^2 f_1 = \begin{bmatrix} \frac{\partial^2 f_1}{\partial w_1^2} & \frac{\partial^2 f_1}{\partial w_1 w_2} \\ \frac{\partial^2 f_1}{\partial w_2 w_1} & \frac{\partial^2 f_1}{\partial w_2^2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$

(c) $\nabla f_1 = (0, 0) \Rightarrow (w_1, w_2) = (0, 0)$

Question 2, f_1

$f_1(w_1, w_2) = \frac{1}{2}w_1^2 + \frac{7}{2}w_2^2$

- (a) Calculate the gradients of the functions. (2pts)
- (b) Are these function convex? (2pts)
- (c) Determine the global minimum of the functions. (2pts)
- (d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

(a) $\nabla f_1 = (\frac{\partial f_1}{\partial w_1}, \frac{\partial f_1}{\partial w_2}) = (w_1, 7w_2)$

(b) $\nabla^2 f_1 = \begin{bmatrix} \frac{\partial^2 f_1}{\partial w_1^2} & \frac{\partial^2 f_1}{\partial w_1 w_2} \\ \frac{\partial^2 f_1}{\partial w_2 w_1} & \frac{\partial^2 f_1}{\partial w_2^2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$

(c) $\nabla f_1 = (0, 0) \Rightarrow (w_1, w_2) = (0, 0)$

(d) $L = 7$ because for x, y
$||(x_1, 7x_2) - (y_1, 7y_2)|| =$
$||(x_1 - y_1, 7x_2 - 7y_2)|| \leqslant 7||(x_1 - y_1, x_2 - y_2)||$
$\Rightarrow \eta = 1/7$

Question 2, f_1

$f_1(w_1, w_2) = \frac{1}{2}w_1^2 + \frac{7}{2}w_2^2$

- (a) Calculate the gradients of the functions. (2pts)
- (b) Are these function convex? (2pts)
- (c) Determine the global minimum of the functions. (2pts)
- (d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

(a) $\nabla f_1 = (\frac{\partial f_1}{\partial w_1}, \frac{\partial f_1}{\partial w_2}) = (w_1, 7w_2)$

(b) $\nabla^2 f_1 = \begin{bmatrix} \frac{\partial^2 f_1}{\partial w_1^2} & \frac{\partial^2 f_1}{\partial w_1 w_2} \\ \frac{\partial^2 f_1}{\partial w_2 w_1} & \frac{\partial^2 f_1}{\partial w_2^2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$

(c) $\nabla f_1 = (0, 0) \Rightarrow (w_1, w_2) = (0, 0)$

(d) $L = 7$ because for x, y
$||(x_1, 7x_2) - (y_1, 7y_2)|| =$
$||(x_1 - y_1, 7x_2 - 7y_2)|| \leqslant 7||(x_1 - y_1, x_2 - y_2)||$
$\Rightarrow \eta = 1/7$

$x_1 = (2, 2)$

Question 2, f_1

$f_1(w_1, w_2) = \frac{1}{2}w_1^2 + \frac{7}{2}w_2^2$

- (a) Calculate the gradients of the functions. (2pts)
- (b) Are these function convex? (2pts)
- (c) Determine the global minimum of the functions. (2pts)
- (d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

(a) $\nabla f_1 = (\frac{\partial f_1}{\partial w_1}, \frac{\partial f_1}{\partial w_2}) = (w_1, 7w_2)$

(b) $\nabla^2 f_1 = \begin{bmatrix} \frac{\partial^2 f_1}{\partial w_1^2} & \frac{\partial^2 f_1}{\partial w_1 w_2} \\ \frac{\partial^2 f_1}{\partial w_2 w_1} & \frac{\partial^2 f_1}{\partial w_2^2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$

(c) $\nabla f_1 = (0,0) \Rightarrow (w_1, w_2) = (0,0)$

(d) $L = 7$ because for x, y
$||(x_1, 7x_2) - (y_1, 7y_2)|| =$
$||(x_1 - y_1, 7x_2 - 7y_2)|| \leqslant 7||(x_1 - y_1, x_2 - y_2)||$
$\Rightarrow \eta = 1/7$

$x_1 = (2, 2)$

$x_2 = (2,2) - \frac{1}{7}(2, 14) = (\frac{12}{7}, 0) \cong (1.7, 0)$

Question 2, f_1

$f_1(w_1, w_2) = \frac{1}{2}w_1^2 + \frac{7}{2}w_2^2$

- (a) Calculate the gradients of the functions. (2pts)
- (b) Are these function convex? (2pts)
- (c) Determine the global minimum of the functions. (2pts)
- (d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

(a) $\nabla f_1 = (\frac{\partial f_1}{\partial w_1}, \frac{\partial f_1}{\partial w_2}) = (w_1, 7w_2)$

(b) $\nabla^2 f_1 = \begin{bmatrix} \frac{\partial^2 f_1}{\partial w_1^2} & \frac{\partial^2 f_1}{\partial w_1 w_2} \\ \frac{\partial^2 f_1}{\partial w_2 w_1} & \frac{\partial^2 f_1}{\partial w_2^2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix}$

(c) $\nabla f_1 = (0,0) \Rightarrow (w_1, w_2) = (0,0)$

(d) $L = 7$ because for x, y
$||(x_1, 7x_2) - (y_1, 7y_2)|| =$
$||(x_1 - y_1, 7x_2 - 7y_2)|| \leqslant 7||(x_1 - y_1, x_2 - y_2)||$
$\Rightarrow \eta = 1/7$

$x_1 = (2,2)$

$x_2 = (2,2) - \frac{1}{7}(2,14) = (\frac{12}{7}, 0) \cong (1.7, 0)$

Obs. $D = 5$, thus for a given ϵ, we will only require $T = \mathcal{O}(\frac{7\times 5^2}{\epsilon})$ steps to get a point close to the minimizer $(0,0)$.

Question 2, f_2

$f_2(w_1, w_2) = 100(w_2 - w_1^2)^2 + (1 - w_1)^2$ (Rosenbrock's function)

a. Calculate the gradients of the functions. (2pts)
b. Are these function convex? (2pts)
c. Determine the global minimum of the functions. (2pts)
d. Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

Question 2, f_2

$f_2(w_1, w_2) = 100(w_2 - w_1^2)^2 + (1 - w_1)^2$ (Rosenbrock's function)

- (a) Calculate the gradients of the functions. (2pts)
- (b) Are these function convex? (2pts)
- (c) Determine the global minimum of the functions. (2pts)
- (d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

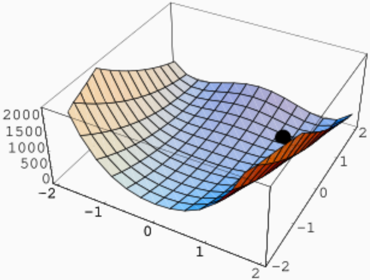
(a) $\nabla f_2 = \left(-400(w_2 - w_1^2)w_1 - 2(1 - w_1), 200(w_2 - w_1^2)\right)$

Question 2, f_2

$f_2(w_1, w_2) = 100(w_2 - w_1^2)^2 + (1 - w_1)^2$ (Rosenbrock's function)

- (a) Calculate the gradients of the functions. (2pts)
- (b) Are these function convex? (2pts)
- (c) Determine the global minimum of the functions. (2pts)
- (d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

(a) $\nabla f_2 = \left(-400(w_2 - w_1^2)w_1 - 2(1 - w_1), 200(w_2 - w_1^2)\right)$



(b)

https://mathworld.wolfram.com/RosenbrockFunction.html

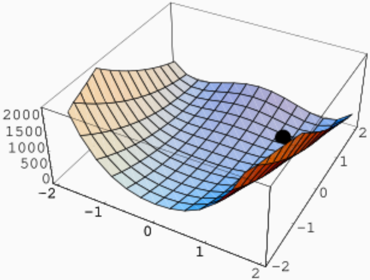
$$\nabla^2 f_2 = \begin{bmatrix} 1200w_1^2 - 400w_2 + 2 & -400w_1 \\ -400w_1 & 200 \end{bmatrix}$$

Question 2, f_2

$f_2(w_1, w_2) = 100(w_2 - w_1^2)^2 + (1 - w_1)^2$ (Rosenbrock's function)

- a) Calculate the gradients of the functions. (2pts)
- b) Are these function convex? (2pts)
- c) Determine the global minimum of the functions. (2pts)
- d) Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

a) $\nabla f_2 = (-400(w_2 - w_1^2)w_1 - 2(1 - w_1), 200(w_2 - w_1^2))$



b)

https://mathworld.wolfram.com/RosenbrockFunction.html

$$\nabla^2 f_2 = \begin{bmatrix} 1200w_1^2 - 400w_2 + 2 & -400w_1 \\ -400w_1 & 200 \end{bmatrix}$$

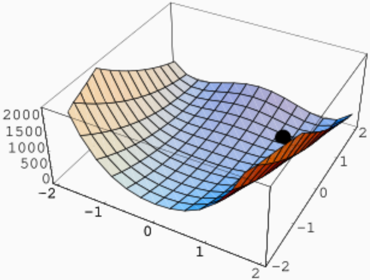
c) $\nabla f_1 = (0, 0) \Rightarrow (w_1, w_2) = (1, 1)$

Question 2, f_2

$f_2(w_1, w_2) = 100(w_2 - w_1^2)^2 + (1 - w_1)^2$ (Rosenbrock's function)

a. Calculate the gradients of the functions. (2pts)
b. Are these function convex? (2pts)
c. Determine the global minimum of the functions. (2pts)
d. Choose a starting point $w = (w_1, w_2)$ within distance 5 from an optimal solution, and perform one step of the Gradient descent algorithm. (2pts)

a. $\nabla f_2 = (-400(w_2 - w_1^2)w_1 - 2(1 - w_1), 200(w_2 - w_1^2))$



b. https://mathworld.wolfram.com/RosenbrockFunction.html

$$\nabla^2 f_2 = \begin{bmatrix} 1200w_1^2 - 400w_2 + 2 & -400w_1 \\ -400w_1 & 200 \end{bmatrix}$$

c. $\nabla f_1 = (0, 0) \Rightarrow (w_1, w_2) = (1, 1)$

d. $L = 400 \Rightarrow \eta = 1/400$

$x_1 = (2, 2)$

$x_2 = (2, 2) - \frac{1}{400}(1602, -400) = (-2.005, 3)$

Question 3

Given a convex, differentiable function $F : K \to \mathbb{R}$ over a convex subset K of $\mathbb{R}^n$, the Bergman divergence of $x, y \in K$ is defined as

$$D_F(x, y) = F(x) - F(y) - \langle \nabla F(y), x - y \rangle$$

Prove that $D_F(x, y) \geqslant 0$

Question 3

Given a convex, differentiable function $F : K \to \mathbb{R}$ over a convex subset K of $\mathbb{R}^n$, the Bergman divergence of $x, y \in K$ is defined as

$$D_F(x, y) = F(x) - F(y) - \langle \nabla F(y), x - y \rangle$$

Prove that $D_F(x, y) \geqslant 0$

Since F is convex, we know by the first order condition that for $x, y \in K$

Question 3

Given a convex, differentiable function $F : K \to \mathbb{R}$ over a convex subset K of $\mathbb{R}^n$, the Bergman divergence of $x, y \in K$ is defined as

$$D_F(x, y) = F(x) - F(y) - \langle \nabla F(y), x - y \rangle$$

Prove that $D_F(x, y) \geqslant 0$

Since F is convex, we know by the first order condition that for $x, y \in K$

$$F(x) \geqslant F(y) + \langle \nabla F(y), x - y \rangle$$

Question 3

Given a convex, differentiable function $F : K \to \mathbb{R}$ over a convex subset K of $\mathbb{R}^n$, the Bergman divergence of $x, y \in K$ is defined as

$$D_F(x, y) = F(x) - F(y) - \langle \nabla F(y), x - y \rangle$$

Prove that $D_F(x, y) \geqslant 0$

Since F is convex, we know by the first order condition that for $x, y \in K$

$$F(x) \geqslant F(y) + \langle \nabla F(y), x - y \rangle$$

$$\Rightarrow F(x) - F(y) - \langle \nabla F(y), x - y \rangle \geqslant 0$$