CSIT 5500 Advanced Algorithms

2020 Spring Semester

Written Assignment 5 solution

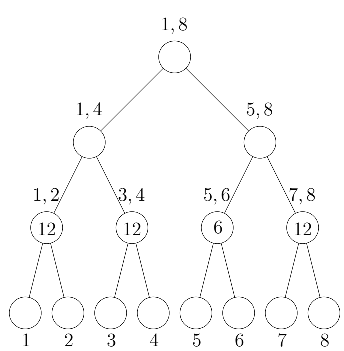
1. (10 points)
   The final array C:

	0	1	2
0	13	14	15
1	13	15	14

a	$\hat{f}_a$
0	14
1	13
2	15
3	14
4	13
5	15
6	14
7	13
8	15
9	14

2. (10 points)

   (a)



   (b) We stop at entry $[3, 4]$ because $12 + 12 > 20$, and we report 4.

3. (10 points)

   - $\nabla f = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$.
     $\langle \nabla f(x_t), z \rangle = 2x_{t_x} z_x + 2x_{t_y} z_y$.
   - D is the convex hull of P, so $\text{argmin}_{z \in D} \langle \nabla f(x_t), z \rangle$ can be obtained by examining all vertices of P.

- $$\begin{aligned} & -\langle \nabla f(x_t), z_t - x_t \rangle \\ = & -2[x_{t_x}(z_{t_x} - x_{t_x}) + x_{t_y}(z_{t_y} - x_{t_y})] \\ = & -2(x_{t_x} z_{t_x} - x_{t_x}^2 + x_{t_y} z_{t_y} - x_{t_y}^2) \\ = & \ 2(x_{t_x}^2 - x_{t_x} z_{t_x} - x_{t_y} z_{t_y} + x_{t_y}^2). \end{aligned}$$

  Stopping condition is $\langle \nabla f(x_t), z_t - x_t \rangle \geq 0$ or $|\langle \nabla f(x_t), z_t - x_t \rangle| \leq \epsilon f(x_t)/(1+\epsilon)$. Below is the proof.

  Let p^* denote the point in D so that $f(p_x^*, p_y^*)$ is the minimum.
  By convexity of f, we have $f(p^*) \geq f(x_t) + \langle \nabla f(x_t), p^* - x_t \rangle$.
  By Lemma 2 in the lecture notes, we have $f(p^*) \geq f(x_t) + \langle \nabla f(x_t), z_t - x_t \rangle$.
  If $\langle \nabla f(x_t), z_t - x_t \rangle \geq 0$, the only possibility is $\langle \nabla f(x_t), z_t - x_t \rangle = 0$, and thus $f(p^*) = f(x_t)$.
  If $\langle \nabla f(x_t), z_t - x_t \rangle < 0$, then $\langle \nabla f(x_t), z_t - x_t \rangle \geq -\epsilon f(x_t)/(1+\epsilon)$ when the algorithm stops. We have

  $$\begin{aligned} & f(p^*) \geq f(x_t) + \langle \nabla f(x_t), z_t - x_t \rangle \\ \rightarrow & f(p^*) \geq f(x_t) - \epsilon f(x_t)/(1+\epsilon) \\ \rightarrow & f(p^*) \geq f(x_t)/(1+\epsilon) \\ \rightarrow & (1+\epsilon) f(p^*) \geq f(x_t) \end{aligned}$$

- $f(x_t + \rho(z_t - x_t)) = (x_{t_x} + \rho(z_{t_x} - x_{t_x}))^2 + (x_{t_y} + \rho(z_{t_y} - x_{t_y}))^2$
  Take the derivative of it with respect to ρ, we have

  $$\begin{aligned} & 2(x_{t_x} + \rho(z_{t_x} - x_{t_x}))(z_{t_x} - x_{t_x}) + 2(x_{t_y} + \rho(z_{t_y} - x_{t_y}))(z_{t_y} - x_{t_y}) \\ = & \ 2x_{t_x}(z_{t_x} - x_{t_x}) + 2\rho(z_{t_x} - x_{t_x})^2 + 2x_{t_y}(z_{t_y} - x_{t_y}) + 2\rho(z_{t_y} - x_{t_y})^2 \\ = & \ 2\rho((z_{t_x} - x_{t_x})^2 + (z_{t_y} - x_{t_y})^2) + 2x_{t_x}(z_{t_x} - x_{t_x}) + 2x_{t_y}(z_{t_y} - x_{t_y}) \end{aligned}$$

  By setting $2\rho((z_{t_x} - x_{t_x})^2 + (z_{t_y} - x_{t_y})^2) + 2x_{t_x}(z_{t_x} - x_{t_x}) + 2x_{t_y}(z_{t_y} - x_{t_y})$ to zero, we have

  $$\begin{aligned} & 2\rho((z_{t_x} - x_{t_x})^2 + (z_{t_y} - x_{t_y})^2) + 2x_{t_x}(z_{t_x} - x_{t_x}) + 2x_{t_y}(z_{t_y} - x_{t_y}) = 0 \\ \rightarrow & \ 2\rho((z_{t_x} - x_{t_x})^2 + (z_{t_y} - x_{t_y})^2) = -2x_{t_x}(z_{t_x} - x_{t_x}) - 2x_{t_y}(z_{t_y} - x_{t_y}) \\ \rightarrow & \ \rho = -(x_{t_x}(z_{t_x} - x_{t_x}) + x_{t_y}(z_{t_y} - x_{t_y}))/((z_{t_x} - x_{t_x})^2 + (z_{t_y} - x_{t_y})^2) \end{aligned}$$

  Let $\rho^* = -\frac{(x_{t_x}(z_{t_x} - x_{t_x}) + x_{t_y}(z_{t_y} - x_{t_y}))}{((z_{t_x} - x_{t_x})^2 + (z_{t_y} - x_{t_y})^2)}$.
  We have

  $$\begin{cases} \operatorname{argmin}_{\rho \in [0,1]} f(x_t + \rho(z_t - x_t)) = \rho^* & \text{if } 0 \leq \rho^* \leq 1. \\ \operatorname{argmin}_{\rho \in [0,1]} f(x_t + \rho(z_t - x_t)) = \operatorname{argmin}_{\rho \in \{0,1\}} f(x_t + \rho(z_t - x_t)) & \text{Otherwise.} \end{cases}$$

- For every pair of points $p, q \in D$,

$$
\begin{aligned}
& ||\nabla f(p) - \nabla f(q)|| \leq L||p - q|| \\
\rightarrow \quad & 2||p - q|| \leq L||p - q|| \\
\rightarrow \quad & 2 \leq L.
\end{aligned}
$$