שאלה 1

נתון: f - L smooth, מקיימת ש' f^* מינימלי.

עדכון:

$$x_t = x_{t-1} - \eta \nabla f(x_{t-1}) \;;\; t \geq 1$$

$$\eta \leq \frac{1}{L}$$

(1) צ"ל:

$$f(x_t) \leq f(x_{t-1}) - \eta\left(1 - \frac{\eta L}{2}\right)\|\nabla f(x_{t-1})\|_2^2$$

הוכחה

כידוע, עבור פונקציות L-smooth,

$$f(y) \leq f(x) + \nabla f(x)^T(y-x) + \frac{L}{2}\|x-y\|_2^2$$

כיוון ש f היא L-smooth, לכן נבחר:

$$\begin{cases} y = x_t = x_{t-1} - \eta \nabla f(x_{t-1}) \\ x = x_{t-1} \end{cases}$$

ונקבל:

$$f(x_t) \leq f(x_{t-1}) + \nabla f(x_{t-1}) \overbrace{(x_t - x_{t-1})}^{-\eta \nabla f(x_{t-1})} + \frac{L}{2}\|x_t - x_{t-1}\|_2^2 =$$

$$= f(x_{t-1}) + \nabla f(x_{t-1})\left(-\eta \nabla f(x_{t-1})\right) + \frac{L}{2}\|-\eta \nabla f(x_{t-1})\|_2^2 =$$

$$= f(x_{t-1}) - \eta\|\nabla f(x_{t-1})\| + \frac{\eta^2 L}{2}\|\nabla f(x_{t-1})\|_2^2 =$$

$$= f(x_{t-1}) - \eta\left(1 - \frac{\eta L}{2}\right)\|\nabla f(x_{t-1})\|_2^2$$

מש"ל

(2) $\eta \le \frac{1}{L}$

צ"ל: $\|\nabla f(x_{t-1})\|_2^2 \le \frac{2}{\eta}\left(f(x_{t-1}) - f(x_t)\right)$

<u>פתרון</u>

נניח אם $\eta \le \frac{1}{L}$ לפי (1)

$$(1) \quad f(x_t) \le f(x_{t-1}) - \eta\left(1 - \frac{\eta L}{2}\right)\|\nabla f(x_{t-1})\|_2^2$$

$$\eta\left(1 - \frac{\eta L}{2}\right)\|\nabla f(x_{t-1})\|_2^2 \le f(x_{t-1}) - f(x_t) \quad / \cdot \frac{2}{\eta}$$

$$2\left(1 - \frac{\eta L}{2}\right)\|\nabla f(x_{t-1})\|_2^2 \le \frac{2}{\eta}\left(f(x_{t-1}) - f(x_t)\right)$$

צ"ל שביטוי הזה גדול/שווה מ-1

$$2 - \eta L \overset{?}{\ge} 1$$

$$1 \ge \eta L$$

$$\frac{1}{L} \overset{?}{\ge} \eta$$

זה אכן כך. מה שאנו מקבלים:

$$\|\nabla f(x_{t-1})\|_2^2 \le \underbrace{2\left(1 - \frac{\eta L}{2}\right)}_{\ge 1}\|\nabla f(x_{t-1})\|_2^2 \le \frac{2}{\eta}\left(f(x_{t-1}) - f(x_t)\right)$$

כלומר:

$$\|\nabla f(x_{t-1})\|_2^2 \le \frac{2}{\eta}\left(f(x_{t-1}) - f(x_t)\right)$$

<u>מש"ל.</u>

(3) צ"ל:

$$\sum_{t=0}^{T} \|\nabla f(x_t)\|_2^2 \le \frac{2}{\eta}\left(f(x_0) - f^*\right)$$

<u>פתרון:</u>

נבצע סכימה על הביטוי מסעיף הקודם.

$$\sum_{t=1}^{T+1} \|\nabla f(x_{t-1})\|_2^2 \le \sum_{t=1}^{T+1} \left[\frac{2}{\eta} f(x_{t-1}) - f(x_t)\right]$$

$$\sum_{t=0}^{T} \|\nabla f(x_t)\|_2^2 \le \frac{2}{\eta} \sum_{t=0}^{T} \left[f(x_t) - f(x_{t+1})\right]$$

נביט באגף ימין:

יש כאן טור טלסקופי; ~~[illegible]~~ נשאר רק עם איבר הראשון והאחרון:

$$\sum_{t=0}^{T} f(x_t) - f(x_{t+1}) = f(x_0) - f(x_{t+1})$$

וכיוון שנתון שהסומה מלמטה ע"י f^*,

כלומר

$$f^* \le f(x_{t+1})$$

לכן

$$f(x_0) - f(x_{t+1}) \le f(x_0) - f^*$$

לכן:

$$\sum_{t=0}^{T} \|\nabla f(x_t)\|_2^2 \le \frac{2}{\eta}\left(f(x_0) - f(x_{t+1})\right) \le \frac{2}{\eta}\left(f(x_0) - f^*\right)$$

מש"ל

(4) נוכיח את הטענה שלמעלה.

$$\sum_{t=0}^{T} \|\nabla f(x_t)\|_2^2 \le \frac{2}{\eta}\left(f(x_0)-f^*\right)$$

אם ניקח את הגרדיאנט המינימלי לאורך כל $t=0\ldots T$
נוכל לרשום:

$$\min_{t=0\ldots T} \|\nabla f(x_t)\|_2^2 \cdot (T+1) \le \underbrace{\sum_{t=0}^{T} \|\nabla f(x_t)\|_2^2 \le \frac{2}{\eta}\left(f(x_0)-f^*\right)}_{\text{מסעיף קודם}}$$

$$\min_{t=0\ldots T} \|\nabla f(x_t)\|_2^2 \le \frac{2}{\eta(T+1)}\left(f(x_0)-f^*\right)$$

$$\min_{t=0\ldots T} \|\nabla f(x_t)\|_2 \le \sqrt{\frac{2}{\eta(T+1)}\left(f(x_0)-f^*\right)}$$

כאשר החלפתי את הסימון t ל- T בשונה
מנוסח התרגיל.

מש"ל

שאלה 2

נתונה $f:\mathbb{R}^n \to \mathbb{R}$ קמורה ו-L-חלקה.

$$x_{t+1} = x_t - \eta_t \nabla f(x_t) \quad , t \geq 0$$

x^* - ממזער

$$h_t = f(x_t) - f(x^*) \quad ; \quad d_t \triangleq \|x_t - x^*\|$$

(1) צ"ל:

$$\forall t \geq 0 : \quad d_{t+1}^2 \leq d_t^2 - 2\eta_t h_t + \eta_t^2 \|\nabla f(x_t)\|_2^2$$

פתרון

$$d_t^2 = \|x_t - x^*\|_2^2 \Rightarrow d_{t+1}^2 = \|x_{t+1} - x^*\|_2^2 =$$

$$= \|(x_{t+1} - x_t) + (x_t - x^*)\|_2^2 =$$

$$= \underbrace{\|x_{t+1} - x_t\|_2^2} + 2\underbrace{(x_{t+1} - x_t)}_{= -\eta \nabla f(x_t)} \overbrace{(x_t - x^*)}^{d_t} + \overbrace{\|x_t - x^*\|_2^2}^{d_t^2}$$

$$\|x_{t+1} - x_t\|_2^2 = \|-\eta_t \nabla f(x_t)\|_2^2$$

$$d_{t+1}^2 = \|-\eta_t \nabla f(x_t)\|_2^2 - 2\eta\, d_t \nabla f(x_t) + d_t^2$$

$$d_{t+1}^2 = d_t^2 - 2\eta\, d_t \nabla f(x_t) + \eta_t^2 \|\nabla f(x_t)\|_2^2 \quad (a)$$

נעזר בתכונות קמורות של הפונקציה: (אי שוויון הגרדיאנט)

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

כאשר נקח: $y = x^*, \ x = x_t$

$$f(x^*) - f(x_t) \geq \nabla f(x_t)(x^* - x_t) \quad / \cdot (-1)$$

$$f(x_t) - f(x^*) \leq \nabla f(x_t)(x_t - x^*)$$

$$h_t \leq \nabla f(x_t)\, d_t \quad (b)$$

נציב את (β) ב-(α) ונקבל את מה שמבוקש:

$$d_{t+1}^2 \le d_t^2 - 2\eta_t h_t + \eta_t^2 \|\nabla f(x_t)\|_2^2$$

מש"ל

(2) גודל צעד דינמי:

$$\eta_t = \frac{h_t}{\|\nabla f(x_t)\|_2^2}$$

פתרון: נציב את גודל הצעד בביטוי מסעיף הקודם:

$$d_{t+1}^2 - d_t^2 \le -\frac{2h_t^2}{\|\nabla f(x_t)\|_2^2} + \frac{h_t^2}{(\|\nabla f(x_t)\|_2^2)^{\cancel{2}}}\cancel{\|\nabla f(x_t)\|_2^2}$$

$$d_{t+1}^2 - d_t^2 \le \frac{-h_t^2}{\|\nabla f(x_t)\|_2^2}$$

נעזר בלמה הנתונה לנו:

$$\|\nabla f(x_t)\|_2^2 \le 2L(f(x_t) - f(x^*)) = 2L h_t$$

ונקבל:

$$d_{t+1}^2 - d_t^2 \le \frac{-h_t^2}{\|\nabla f(x_t)\|_2^2} \le \frac{-h_t^2}{2Lh_t} = -\frac{h_t}{2L}$$

$$d_{t+1}^2 - d_t^2 \le \frac{h_t}{2L}$$

מש"ל

(3) צ"ל: לכל $T \ge 1$ מתקיים:

$$\sum_{t=0}^{T} h_t \le 2L d_0^2$$

<u>פתרון:</u>

מהלמה מסעיף הקודם, עבור איטרציה:

$$d_{t+1}^2 - d_t^2 \le -\frac{h_t}{2L}$$

$$h_t \le -2L\left(d_{t+1}^2 - d_t^2\right) = 2L\left(d_t^2 - d_{t+1}^2\right)$$

נבצע סכום:

$$\sum_{t=0}^{T-1} h_t \le \sum_{t=0}^{T-1} 2L\left(d_t^2 - d_{t+1}^2\right)$$

יש כאן טור טלסקופי:

$$\sum_{t=0}^{T-1} h_t \le 2L\left(d_0^2 - \underbrace{d_T^2}_{\ge 0}\right) \le 2L d_0^2$$

$$\sum_{t=0}^{T-1} h_t \le 2L d_0^2 \quad (C)$$

<u>מסקנה</u>

יש להוכיח:

$$f(\bar{x}_T) - f(x^*) \le \frac{2L d_0^2}{T} = \frac{2L\|x_0 - x^*\|_2^2}{T}$$

<u>פתרון</u>

נציב h_t לתוך ביטוי (C)

$$\sum_{t=0}^{T-1} f(x_t) - f(x^*) \le 2L d_0^2 \quad / \cdot \frac{1}{T}$$

$$\frac{1}{T}\sum_{t=0}^{T-1} f(x_t) - f(x^*) \le \frac{2L d_0^2}{T}$$

נשתמש באי שוויון Jensen:

$$f\left(\sum_{i=1}^{N} \lambda_i x_i\right) \le \sum_{i=1}^{N} \lambda_i f(x_i) \quad ; \quad \sum \lambda_i = 1$$

ולכן:

$$\bar{x}_T = \frac{1}{T}\sum_{t=0}^{T-1} x_t$$

$$f(\bar{x}_T) \le \frac{1}{T}\sum_{t=0}^{T-1} f(x_t)$$

כאשר $\lambda_{0 \dots T-1} = \frac{1}{T}$

לכן נקבל:

$$f(\bar{x}_T) - f(x^*) \le \frac{1}{T}\sum_{t=0}^{T-1} f(x_t) - \underbrace{\frac{1}{T}\sum_{t=0}^{T-1} f(x^*)}_{= f(x^*)}$$

$$f(\bar{x}_T) - f(x^*) \le \frac{1}{T}\sum_{t=0}^{T-1} f(x_t) - f(x^*) \underbrace{\le \frac{2L d_0^2}{T}}_{\text{הכחנו לפני}}$$

ונקבל:

$$f(\bar{x}_T) - f(x^*) \le \frac{2L d_0^2}{T} = \frac{2L \| x_0 - x^* \|^2}{T}$$

מש"ל

שאלה 3 נתונה בעיה הבאה:

$$\max_{0 \le x \le 5} \left\{ 2\sqrt{5-x} + \frac{x}{\sqrt{5}} \right\}$$

(1)

נבדוק את נגזרת של הפונק' בתחום:

$$f(x) = 2\sqrt{5-x} + \frac{x}{\sqrt{5}}$$

$$f'(x) = -\frac{1}{2} \cdot 2 \cdot (5-x)^{-\frac{1}{2}} + \frac{1}{\sqrt{5}} = \frac{1}{\sqrt{5}} - \frac{1}{\sqrt{5-x}} =$$

$$= \frac{\sqrt{5-x} - \sqrt{5}}{\sqrt{5(5-x)}}$$

• מכנה חיובי

• $\sqrt{5} \ge \sqrt{5-x}$ עבור $x \in [0,5]$

לכן המונה שלילי

ולכן $f'(x) < 0 \quad \forall \; x \in [0,5]$

לכן הפונק' יורדת לאורך $0 \le x \le 5$

והמקסימום שלה מתקבל ב $x=0$

$$f(0) = 2\sqrt{5}$$

זהו הפתרון האופטימלי!

לכן לכל $0 \le x \le 5$ מתקיים

$$2\sqrt{5-x} + \frac{x}{\sqrt{5}} \le 2\sqrt{5}$$

(2) צ"ל:

$$\sum_{i=1}^{n} \frac{a_i}{\sqrt{\sum_{j=1}^{i} a_j}} \le 2\sqrt{\sum_{i=1}^{n} a_i}$$

לכל סדרת מס' אי-שליליים
$a_1 \dots a_n$.

קודם נבדוק שמתקיים עבור $n=1$.

$$\frac{a_1}{\sqrt{a_1}} \le 2\sqrt{a_1}$$

$$\sqrt{a_1} \le 2\sqrt{a_1} \quad \checkmark$$

נניח כי הטענה תקפה עבור n טבעי, ונוכיח
שהיא תקפה גם עבור $n+1$.
נצטרך להוכיח:

$$\sum_{i=1}^{n+1} \frac{a_i}{\sqrt{\sum_{j=1}^{i} a_j}} \le 2\sqrt{\sum_{i=1}^{n+1} a_i}$$

$$\sum_{i=1}^{n+1} \frac{a_i}{\sqrt{\sum_{j=1}^{i} a_j}} = \sum_{i=1}^{n} \frac{a_i}{\sqrt{\sum_{j=1}^{i} a_j}} + \frac{a_{n+1}}{\sqrt{\sum_{j=1}^{n+1} a_j}} \le 2\sqrt{\sum_{i=1}^{n} a_i} + \frac{a_{n+1}}{\sqrt{\sum_{j=1}^{n+1} a_j}}$$

נסמן:

$$\left\{ \begin{array}{l} x = a_{n+1} \\ S = \sum_{i=1}^{n+1} a_i \end{array} \right. \Leftarrow \underline{0 \le x \le S}$$

כי $a_1 \dots a_n \ge 0$

כלומר $S - x = \sum_{i=1}^{n} a_i$

ונקבל:

$$\sum_{i=1}^{n+1} \frac{a_i}{\sqrt{\sum_{j=1}^{i} a_j}} \le 2\sqrt{S-x} + \frac{x}{\sqrt{S}} \le 2\sqrt{S} = 2\sqrt{\sum_{i=1}^{n+1} a_i}$$

מסעיף הקודם

כלומר, הוכחנו את הטענה.

מש"ל