

-ל גורם

+ 3. נסחף נסחף ועומק ה- $\nabla f(x)$ + פונקציית

$$x_t = x_{t-1} - \eta \nabla f(x_{t-1}) \quad \text{האריך ורץ}$$

$$\underline{\eta \leq 1/L}$$

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

: L-smooth מבחן גוף כינון

$$f(x_t) \leq f(x_{t-1}) + \nabla f(x_{t-1})^\top (x_t - x_{t-1}) + \frac{\ell}{2} \|x_t - x_{t-1}\|_2^2$$

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

$$\Rightarrow f(x_t) \leq f(x_{t-1}) + \nabla f(x_{t-1})^\top (-\eta \nabla f(x_{t-1})) + \frac{\ell}{2} \eta^2 \|\nabla f(x_t)\|_2^2$$

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

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

$$\frac{2}{\eta} (f(x_{t-1}) - f(x_t)) \geq \|\nabla f(x_{t-1})\|_2^2 \quad \text{ס. 2}$$

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

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

$$\Leftrightarrow \frac{\ell}{2} \eta \leq \eta \left(\ell - \frac{\ell}{2} \eta\right)$$

$$\Leftrightarrow \frac{1}{2} \leq 1 - \frac{\ell}{2} \eta$$

$$\Leftrightarrow \eta \leq 1/\ell$$

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

✓

■

$$\sum_{t=0}^T \|\nabla f(x_t)\|_2^2 \leq \frac{2}{\eta} (f(x_0) - f^*) \quad \text{3.3}$$

$$\|\nabla f(x_t)\|_2^2 \leq \frac{2}{\eta} (f(x_t) - f(x_{t+1})) \quad : 2.7.60N$$

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

$$= \frac{2}{\eta} [f(x_0) - f(x_1) + f(x_1) - f(x_2) + \dots + f(x_T) - f(x_{T+1})] \quad \text{2.7.60C 7.16}$$

$$= \frac{2}{\eta} [f(x_0) - f(x_{T+1})] \leq \frac{2}{\eta} (f(x_0) - f^*) \quad \checkmark$$

$f^* \leq f(x_{T+1})$

$$\sum_{t=0}^T \|\nabla f(x_t)\|_2^2 \leq \frac{2}{\eta} (f(x_0) - f^*) \quad \blacksquare$$

$$\min_{T=0, \dots, t} \|\nabla f(x_T)\|_2 \leq \sqrt{\eta \frac{2}{(t+1)} (f(x_0) - f^*)} \quad .4$$

$$\frac{\ell}{t+1} \cdot \sum_{T=0}^t \|\nabla f(x_T)\|_2^2 \leq \frac{2}{\eta(t+1)} (f(x_0) - f^*) \quad \frac{3.8.00N}{\blacksquare}$$

$$\min_{T=0, \dots, t} \|\nabla f(x_T)\|_2^2 \leq \mathbb{E} [\|\nabla f(x_T)\|_2^2] \leq \frac{2}{\eta(t+1)} (f(x_0) - f^*) \quad \checkmark$$

ההנחתה מושגית היא $\|\nabla f(x_T)\|_2^2 \geq 0$ ומכיוון ש- \mathbb{E} מוגדרת כ- $\mathbb{E}[f]$, אז $\mathbb{E}[\|\nabla f(x_T)\|_2^2] \geq \mathbb{E}[\|\nabla f(x_T)\|_2^2]$.

$$\min_{T=0, \dots, t} \|\nabla f(x_T)\|_2 \leq \sqrt{\frac{2}{\eta(t+1)} (f(x_0) - f^*)}$$

-25.873

L-Smooth-1 convex - גמיש ותבוקה פונקציית $f: \mathbb{R}^n \rightarrow \mathbb{R}$ יי

הנובע מושג PD GD פרימיט פונקציית

$$x_{t+1} = x_t - \eta_t \nabla f(x_t) ; \quad \{\eta_t\}_{t=0}^{\infty}$$

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

$$d_{t+1}^2 \leq d_t^2 - 2\eta_t h_t + \eta_t^2 \|\nabla f(x_t)\|_2^2 \stackrel{t \geq 0}{\text{for}} \quad \underbrace{\text{ס. 3.}}_{\text{ל}}$$

לפנינו מושג הוכחה

$$f(x^*) \geq f(x_t) + \nabla f(x_t)^\top (x^* - x_t)$$

$$\Rightarrow \nabla f(x_t)^\top (x_t - x^*) \geq f(x_t) - f(x^*) = h_t$$

✓

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

$$\|x_t - x^*\|_2^2 - 2\eta_t \nabla f(x_t)^\top (x_t - x^*) + \eta_t^2 \|\nabla f(x_t)\|_2^2$$

$$\Rightarrow d_{t+1}^2 = d_t^2 - 2\eta_t \nabla f(x_t)^\top (x_t - x^*) + \eta_t^2 \|\nabla f(x_t)\|_2^2$$

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

□

✓

$$\eta_t = \frac{h_t}{\|\nabla f(x_t)\|_2^2} \quad \text{--- ပုဂ္ဂန် ဒါန် အောင်}$$

ပုဂ္ဂန် ②

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

$$\begin{aligned} d_{t+1}^2 - d_t^2 &\leq -2\eta_t h_t + \eta_t^2 \|\nabla f(x_t)\|_2^2 \\ &= \frac{-2\eta_t h_t^2}{\|\nabla f(x_t)\|_2^2} + \frac{h_t^2}{\|\nabla f(x_t)\|_2^4} \|\nabla f(x_t)\|_2^2 \\ &= \frac{-h_t^2}{\|\nabla f(x_t)\|_2^2} \end{aligned}$$

✓

အသေ

$$* \|\nabla f(x_t)\|_2^2 \leq 2L(f(x_t) - f(x^*)) \leq 2Lh_t$$

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

အိမ် အား ပုဂ္ဂန် ပါ။ ဖြစ် အောင်

အိမ် အား ပုဂ္ဂန်

✓

$$\sum_{t=0}^{T-1} h_t \leq 2L d_0^2$$

PPTN T≥1 68 3

$$h_t \leq 2L(d_t^2 - d_{t+1}^2)$$

: 2.8.6 ON

$$\sum_{t=0}^{T-1} h_t = 2L \sum_{t=0}^{T-1} (d_t^2 - d_{t+1}^2) = 2L [d_0^2 - d_1^2 + d_1^2 - d_2^2 + \dots + d_{T-1}^2 - d_T^2]$$

: 2.10.6 C 710

$$\leq 2L(d_0^2 - d_T^2) \leq 2L d_0^2. \quad \checkmark$$

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

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

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

\uparrow

• 10) ' PPTN 1.1

\checkmark

3. פונק.

(ר' מיל' נס' 10. חישוב נאכ'ת ה-3 נאכ'ת ה-3 נאכ'ת ה-3)

$$\max_{0 \leq x \leq S} \left\{ 2\sqrt{S-x} + \frac{x}{\sqrt{S}} \right\}$$

. 2 \sqrt{S} הינו פון. פון. פון.

כינון:

$$f(x) = 2\sqrt{S-x} + \frac{x}{\sqrt{S}} : 0 \leq x \leq S$$

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

$$= \frac{1}{\sqrt{S}} - \frac{1}{\sqrt{S-x}}$$

: $x \geq 0$ סדרה מוגדרת

$$\frac{1}{\sqrt{S}} - \frac{1}{\sqrt{S-x}} \leq 0$$

$$\Leftrightarrow \frac{1}{\sqrt{S}} \leq \frac{1}{\sqrt{S-x}}$$

$$\Leftrightarrow \sqrt{S-x} \leq \sqrt{S}$$

$$\Leftrightarrow S-x \leq S$$

$$x \geq 0$$

אך כראוי הוכיחו ה-3 נס' ה-3 נס' \Leftrightarrow ה-3 נס'

$x=0$ מוגדר ב-0 נס' נס' נס' נס'

$$\max_{0 \leq x \leq S} f(x) = f(0) = 2\sqrt{S-0} + \frac{0}{\sqrt{S}} = 2\sqrt{S}$$

$$2\sqrt{S-x} + \frac{x}{\sqrt{S}} \leq 2\sqrt{S} \quad \forall 0 \leq x \leq S. \text{ פון.}$$

3.2

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

.ר' פיל a_1, \dots, a_n ר' פיל ג' 65

(! נון כוונתית)

$$\sum_{i=1}^1 \frac{a_i}{\sqrt{\sum_{j=1}^i a_j}} = \frac{a_1}{\sqrt{a_1}} = \sqrt{a_1} \leq 2 \sqrt{a_1}$$

נניח כי הטענה נכונה עבור $N=k$

(נניח כי הטענה נכונה עבור $N=k+1$)

$$\begin{aligned} \sum_{i=1}^{k+1} \frac{a_i}{\sqrt{\sum_{j=1}^i a_j}} &= \sum_{i=1}^k \frac{a_i}{\sqrt{\sum_{j=1}^i a_j}} + \frac{a_{k+1}}{\sqrt{\sum_{j=1}^{k+1} a_j}} \\ &\leq \underbrace{2 \sqrt{\sum_{i=1}^k a_i}}_{\text{הערכות}} + \frac{a_{k+1}}{\sqrt{\sum_{j=1}^{k+1} a_j}} \\ &= 2 \sqrt{\sum_{i=1}^{k+1} a_i - a_{k+1}} + \frac{a_{k+1}}{\sqrt{\sum_{j=1}^{k+1} a_j}} \end{aligned}$$

$$0 \leq a_{k+1} \leq \sum_{i=1}^{k+1} a_i \Rightarrow \exists x | x = a_{k+1} - s \quad s = \sum_{j=1}^{k+1} a_j$$

$$= 2 \sqrt{s-x} + \frac{x}{\sqrt{s}} \leq 2 \sqrt{s} = 2 \sqrt{\sum_{i=1}^{k+1} a_i}$$

ר' פילון ✓

ר' פיל ג' 65 סעיף ב' הוכיח כוונתית כ הטענה וטענה

.ר' פיל a_1, \dots, a_n

$\min_{x \in C} f(x)$ הוכיח כי הפונקציית $\frac{-4\sqrt{2x+4}}{\sqrt{x+3}}$ ולשונו.

לעתם נזכיר $x \in \mathbb{R}^n$ ו- $x^* \in \mathbb{R}^m$ כוונתית $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ נסמן \hat{x} ו- \hat{x}^* .

$$\|\nabla f(x)\|_2^2 \leq 2L(f(x) - f(x^*))$$

הו' מ-ה
בהתאם ל (מי ימיה פותח תקופה) (מי תקונן קולו)
ולפניהם GD גודל צבוי קולך לאירועים:

$$\|\nabla f(x_{t+1})\|_2^2 \leq 2L(f(x_{t+1}) - f(x_t)) \quad \forall t \geq 1$$

$\gamma_{C_1} \rightarrow C_1$

$$\|\nabla f(x_0)\|_2^2 \leq 2L(f(x_0) - f(x_1))$$

$$\leq 2L(f(x_0) - f(x^*))$$

$$\|\nabla f(x)\|_2^2 \leq 2L(f(x) - f(x^*))$$

$$\sum_{t=1}^T (f(x_t) - f(x^*)) \leq 4LD^2$$

: 3 . 2

$$\sum_{t=1}^T (f(x_t) - f(x^*)) \leq \sqrt{2D^2 \sum_{t=1}^T \|g_t\|_2^2}$$

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

. פונקציית ה-grad f

$$\sum_{t=1}^T (f(x_t) - f(x^*)) \leq \sqrt{2D^2 \cdot 2L \sum_{t=1}^T f(x_t) - f(x^*)}$$

$$\frac{\sum_{t=1}^T (f(x_t) - f(x^*))}{\sqrt{\sum_{t=1}^T f(x_t) - f(x^*)}} \leq \sqrt{4LD^2}$$

$$\sqrt{\sum_{t=1}^T (f(x_t) - f(x^*))} \leq \sqrt{4LD^2} / \sqrt{2} \quad . \text{טביה וטב}$$

$$\sum_{t=1}^T f(x_t) - f(x^*) \leq 4LD^2$$

$$f(\bar{x}_T) - f(x^*) \leq \frac{1}{T} \sum_{t=1}^T f(x_t) - f(x^*) \leq \frac{4LD^2}{T}$$

ט) ||||\ell||.

✓

-5 סעיפים

1. מינימיזציה:

$$\min_x \bar{F}(x) = \frac{\ell}{m} \sum_{i=1}^m f_i(x)$$

L-smooth $\Rightarrow \{\bar{f}_i\}_{i=1}^m$

$$g = \nabla f_i(x) \quad i \sim \text{Uniform}\{1, \dots, m\}$$

$$x_{t+1} = x_t - \eta g_t$$

$$h_t = F(x_t) - F(x^*) \quad ; \quad \eta = \frac{\mu}{L^2} = \frac{n}{L} \quad ; \quad \mu = \frac{\mu}{L} \quad \frac{100}{1}$$

מגדירים $F(x)$ סעיף 3.0

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

$$\sum_{i=1}^m f_i(y) \leq \sum_{i=1}^m f_i(x) + \sum_{i=1}^m \nabla f_i(x)^T (y - x) + m \frac{L}{2} \|y - x\|_2^2$$

$$\frac{\ell}{m} \sum_{i=1}^m f_i(y) \leq \frac{\ell}{m} \sum_{i=1}^m f_i(x) + \frac{\ell}{m} \sum_{i=1}^m \nabla f_i(x)^T (y - x) + \frac{\ell}{2} \|y - x\|_2^2$$

$$\bar{F}(y) \leq \bar{F}(x) + \nabla \bar{F}(x)^T (y - x) + \frac{\ell}{2} \|y - x\|_2^2$$

מגדירים סעיף 1.0 F פולינומיאלי



$$h_{t+1} - h_t \leq -\eta \nabla F(x_t)^\top g_t + \frac{\gamma^2}{2} \|g_t\|_2^2 \quad \text{f3.2}$$

-גראן
ט-גראן גראן

$$F(x_{t+1}) - F(x_t) \leq \nabla F(x_t)^\top (x_{t+1} - x_t) + \frac{\gamma^2}{2} \|x_{t+1} - x_t\|_2^2$$

$$x_{t+1} - x_t = -\eta g_t$$

: יונס-ט-גראן

$$F(x_{t+1}) - F(x^*) - (F(x_t) - F(x^*)) \leq \nabla F(x_t)^\top (-\eta g_t) + \frac{\gamma^2}{2} \|g_t\|_2^2$$

$$h_{t+1} - h_t \leq -\eta \nabla F(x_t)^\top g_t + \frac{\gamma^2}{2} \|g_t\|_2^2 \quad \checkmark$$

$$\mathbb{E}[h_{t+1} - h_t] \leq -\eta^2 \mathbb{E}[h_t] \quad \text{f3.3}$$

$$\mathbb{E}[h_{t+1} - h_t] \leq \mathbb{E}[-\eta \nabla F(x_t)^\top g_t] + \frac{\gamma^2}{2} \eta^2 \mathbb{E}\|g_t\|_2^2$$

: פונקציית אפקט שלוניה

$$\mathbb{E}\|g_t\|_2^2 = \mathbb{E}[\mathbb{E}\|g_t\|_2^2 | x_t] \leq \mathbb{E}[2L(F(x_t) - F(x^*))]$$

$$\leq 2L \mathbb{E}[h_t] \quad \checkmark$$

$$\mathbb{E}[-\eta \nabla F(x_t)^\top g_t] = \mathbb{E}[\mathbb{E}[-\eta \nabla F(x_t)^\top g_t | x_t]]$$

$$= \mathbb{E}[-\eta \nabla F(x_t)^\top \mathbb{E}[g_t | x_t]] = \mathbb{E}[-\eta \nabla F(x_t)^\top \nabla F(x_t)]$$

↑
... דוגמאות רגולריזציה .

✓

$$= -\eta \mathbb{E}\|\nabla F(x_t)\|_2^2 \leq -\eta 2\mu \mathbb{E}[F(x_t) - F(x^*)] \leq -2\mu\eta \mathbb{E}[h_t]$$

הוכחה של הטענה

$$\mathbb{E}[h_{t+1} - h_t] \leq -2\mu\gamma \mathbb{E}[h_t] + \frac{\gamma^2}{2} \cdot 2L \mathbb{E}[h_t]$$

$$= \left(-2\mu \cdot \frac{\mu}{L^2} + \gamma^2 \cdot \frac{\mu^2}{L^2} \right) \mathbb{E}[h_t] = (-2\mu^2 + \mu^2) \mathbb{E}[h_t]$$

$$\Rightarrow \mathbb{E}[h_{t+1} - h_t] \leq -\mu^2 \mathbb{E}[h_t] \quad \checkmark$$

$$\mathbb{E}[h_t] \leq e^{-\mu^2 t} \mathbb{E}[h_0] \quad \text{ג 4}$$

הוכחה:

$$\mathbb{E}[h_{t+1}] - \mathbb{E}[h_t] \leq -\mu^2 \mathbb{E}[h_t]$$

'3 Proof

$$\mathbb{E}[h_{t+1}] \leq (\lambda - \mu^2) \mathbb{E}[h_t] \quad \forall t \geq 0$$

1)

$$\mathbb{E}[h_t] \leq (\lambda - \mu^2) \mathbb{E}[h_{t-1}] \leq (\lambda - \mu^2)^2 \mathbb{E}[h_{t-2}] \leq \dots \leq (\lambda - \mu^2)^t \mathbb{E}[h_0]$$

$$\forall x, y \in \mathbb{R}^n \quad \text{כו ס. פ. ל.}$$

$$F(y) \leq F(x) + \nabla F(x)^\top (y-x) + \frac{\mu}{2} \|y-x\|_2^2 \quad \begin{matrix} \text{smooth} \\ \text{2-s} \end{matrix}$$

$$F(y) \geq F(x) + \nabla F(x)^\top (y-x) + \frac{\mu}{2} \|y-x\|_2^2 \quad \begin{matrix} \text{Strongly} \\ \text{Convex} \end{matrix}$$

$$\Rightarrow F(x) + \nabla F(x)^\top (y-x) + \frac{\mu}{2} \|y-x\|_2^2 \leq F(x) + \nabla F(x)^\top (y-x) + \frac{\mu}{2} \|y-x\|_2^2$$

$$\Rightarrow \frac{\mu}{2} \leq \frac{\mu}{2} \Rightarrow \frac{\mu}{L} \leq 1 \Rightarrow \boxed{\mu \leq L}$$

$$0 < r \leq 1 \quad \text{for} \quad e^{-r^2} \geq (1-r^2) \quad \text{ר'ז}$$

$$\omega(r) = e^{-r^2} - (1-r^2)$$

$$\omega'(r) = -2re^{-r^2} + 2r$$

$$\Leftrightarrow r > 0 \quad e^{-r^2} \leq 1$$

$$\Leftrightarrow 1 \leq e^{r^2}$$

$$\Leftrightarrow 0 < r$$

לפי הדרישה, מינימום הערך של $\omega(r)$ מושג ב-

$$\omega(0) = 1 - (1-0) = 0$$

נוכיח כי $\omega(r)$ מונוטונית י递增 ב-

$$\omega(r) = e^{-r^2} - (1-r^2) \geq 0 \quad \omega(0) = 0 \quad \text{ר'ז}$$

$$\mathbb{E}[h_t] \leq (1-r^2)^t \mathbb{E}[h_0] \leq e^{-r^2 t} \mathbb{E}[h_0]$$

