IML- Ex.6

Bar Aloni- 305230740

PCA:

1. 😊
2. 😊

Convex optimization and SGD:

3.

a) יהיו $f_i: \mathcal{V} \to \mathbb{R}$ ל-$i \in [m]$ פונקציות קמורות.

נוכיח ש: $g: \mathcal{V} \to \mathbb{R}$ המוגדרת ע"י $g(u) = \Sigma_{i=1}^{m} \gamma_i f_i(u)$ קמורה:

מקמירות f_i לכל i, מתקיים לכל $u, v \in \mathcal{V}, \alpha \in [0,1]$:

$f_i(\alpha u + (1-\alpha)v) \leq \alpha f_i(u) + (1-\alpha) f_i v$

ולכן:

$$
\begin{aligned}
g(\alpha u + (1-\alpha)v) &= \Sigma_{i=1}^{m} \gamma_i [\, f_i(\alpha u + (1-\alpha)v)] \leq \\
&\leq \Sigma_{i=1}^{m} \gamma_i [\alpha f_i(u) + (1-\alpha) f_i v] = \\
&= \Sigma_{i=1}^{m} \alpha \gamma_i f_i(u) + (1-\alpha)\gamma_i f_i v = \\
&= \alpha \Sigma_{i=1}^{m} \gamma_i f_i(u) + (1-\alpha) \Sigma_{i=1}^{m} \gamma_i f_i v = \\
&= \alpha g(u) + (1-\alpha) g v
\end{aligned}
$$

b) יהיו $f, g, h: \mathbb{R} \to \mathbb{R}$ המוגדרות ע"י: $f(x) = -x,\ g(x) = |x|$.

נבחין ש: f,g קמורות, אך: $h(x) = f(g(x)) = -|x|$ איננה קמורה.

c) תהא $f: C \to \mathbb{R}$, C קמורה.

נוכיח ש: f קמורה אמ"מ $epi(f) = \{(u,t) | f(u) \leq t\}$. יהיו $u, v \in C, s, t \in \mathbb{R}, \alpha \in [0,1]$. מתקיים:

$\Leftarrow$:

מקמירות C: $\alpha u + (1-\alpha)v \in C$ ולכן: $f(\alpha u + (1-\alpha)v)$ מוגדר ומקיים:

$f(\alpha u + (1-\alpha)v) \leq \alpha f(u) + (1-\alpha) f(v)$.

מכיוון ש:$u, v \in C$ אז קיימים $s, t \in \mathbb{R}$ כך ש: $(u,t), (v,s) \in epi(f)$.

כלומר:$f(u) \leq t, f(v) \leq s$ ולכן: $\alpha f(u) + (1-\alpha) f(v) \leq \alpha t + (1-\alpha)s$

וקיבלנו ש: $(\alpha u + (1-\alpha)v, \alpha t + (1-\alpha)s) \in epi(f)$.

נבחין ש:$(\alpha u + (1-\alpha)v, \alpha t + (1-\alpha)s) = \alpha(u,t) + (1-\alpha)(v,s)$

כלומר $epi(f)$ אכן קבוצה קמורה כנדרש.

$\Rightarrow$:

$epi(f)$ קבוצה קמורה, ויהיו $u, v \in C, s, t \in \mathbb{R}, \alpha \in [0,1]$ כך ש:

$(u,t), (v,s) \in epi(f)$. ו-u, v הן נקודות גבול של האפיגרף, כלומר:

$f(u) = t, f(v) = s$ אזי:

$\alpha(u,t) + (1-\alpha)(v,s) = (\alpha u + (1-\alpha)v, \alpha t + (1-\alpha)s) \in epi(f)$

כלומר: $f(\alpha u + (1-\alpha)v) \leq \alpha t + (1-\alpha)s = \alpha f(u) + (1-\alpha) f(v)$.

ו-f היא אכן פונקציה קמורה כנדרש.

d) יהיו $f_i: \mathcal{V} \to \mathbb{R}$ ל-$i \in I$ פונקציות קמורות.

נוכיח ש: $f: \mathcal{V} \to \mathbb{R}$ המוגדרת ע"י $f(u) = \sup_{i \in I} f_i(u)$ קמורה:

נתבונן על האפיגרף של f:

$$epi(f)=\{(u,t)|f(u)\le t\}=\left\{(u,t)\middle|\sup_{i\in I}f(u)\le t\right\}$$
$$=\{(u,t)|\forall i\in I:\ f_i(u)\le t\}=\ \cap_{i\in I}\{(u,t)|f_i(u)\le t\}$$

ומכיוון שלכל $i \in I$: $\{(u,t)|f_i(u)\le t\}$ היא קבוצה קמורה כאפיגרף של הפונקציה הקמורה f_i (מסעיף קודם), ומהטענה שהוכחנו בתרגול 13 לפיה חיתוך קובצות קמורה היא קבוצה קמורה, נקבל ש: $epi(f)$ היא קבוצה קמורה, כלומר, גם הפונקציה f קמורה (מסעיף קודם).

.4

a) <u>נוכיח ש:</u> $l_{x,y}^{hinge}(w,b)=\max\,(0,1-y(\langle w,x\rangle+b))$ עבור $x\in\mathbb{R}^d$, $y\in\{\pm1\}$ היא פונקציה קמורה:

נבחין שלכל y נתון, הארגומנט $1-y(\langle w,x\rangle+b))$ הוא פונקציה אפינית ולכן קמורה (כפי שהוכחנו בתרגול 13), וכי הפונקציה 0 היא פונקציה קבועה ולכן קמורה אף היא._מתרגיל 3.d הראינו שסופרימום של פונקציות קמורות היא פונקציה קמורה, ולכן מקסימום (סופרימום ששיך לקבוצה) על פונקציות קמורות היא פונקציה קמורה. כלומר: $l_{x,y}^{hinge}$ היא פונקציה קמורה כנדרש.

b) מטענה שהראינו בתרגול מתקיים שאם $f=\max_{i\in I}f_i:\mathcal{V}\to\mathbb{R}$ עבור $\{f_i:\mathcal{V}\to\mathbb{R}\}_{i\in I}$ פונקציות קמורות, אזי עבור $u\in\mathcal{V}$ מתקיים: $\partial f_j(u)\subset\ \partial f(u)$ ל- $j=x_{i\in I}f_i(u)$.

נחשב את $\partial f_i(w,b)$ ל-$i\in[2]$ של הפונקציה:
$l_{x,y}^{hinge}(w,b)=\max\,(0,1-y(\langle w,x\rangle+b))$, נגזור לפי w.
מכיוון ש-2 הארגומנטים דיפרנציאבילים, ממשפט אחר שהוכחנו בתרגול נובע ש:

$$\partial f_1(w,b)=\{\Delta f_1(w,b)\}=\{(0'(w,b)\}=\{0\}$$
$$\partial f_2(w,b)=\{\Delta f_2(w,b)\}=\{(1-y(\langle w,x\rangle+b))'(w,b)\}=\{-yx\}$$

ולכן:$g\in\partial f(w,b)$ עבור-

$$g(w,b)=\begin{cases}0 & if\ 1-y(\langle w,x\rangle+b)\le 0\\ -yx & else\end{cases}$$

c) יהיו $f_i:\mathbb{R}^d\to\mathbb{R}$ ל-$i\in[m]$ פונקציות קמורות, ונגדיר: $f=\Sigma_{i=1}^m f_i:\mathbb{R}^d\to\mathbb{R}$.
יהיו $\gamma_i\in\partial f_i(x)$ לכל $i\in[m]$.
<u>נוכיח ש:</u> $\Sigma_{i\in[m]}\,\gamma_i\in\partial\Sigma_{i\in[m]}\,f_i(x)$:
מכך ש לכל $i\in[m]$: $\gamma_i\in\partial f_i(x)$, אז לכל $y\in\mathbb{R}^d$:
$f_i(y)\ge f_i(x)+\langle\gamma_i,y-x\rangle$. נסכום על i:

$$f(y)=\ \Sigma_{i\in[m]}f_i(y)\ge\Sigma_{i\in[m]}(f_i(x)+\langle\gamma_i,y-x\rangle)$$
$$=\Sigma_{i\in[m]}f_i(x)+\Sigma_{i\in[m]}\langle\gamma_i,y-x\rangle$$
$$=\ f(x)+\ \left\langle\Sigma_{i\in[m]}\gamma_i,y-x\right\rangle$$

מכיוון ש: x,y לא תלויים ב-i.
אז קיבלנו ש: $\Sigma_{i\in[m]}\,\gamma_i\in\partial\Sigma_{i\in[m]}\,f_i(x)$ כנדרש.

d) עבור $\lambda\ge0$, $\{x_i,y_i\}_{i=1}^m\subset\mathbb{R}^d\ \times\{\pm1\}$ נגדיר: $f:\mathbb{R}^d\ \times\mathbb{R}\to\mathbb{R}$ ע"י: $f(w,b)=\frac{1}{m}\Sigma_{i=1}^m l_{x_i,y_i}^{hinge}(w,b)+\frac{\lambda}{2}||w||^2$. מתקיים:

i. f קמורה:

$l_{x_i,y_i}^{hinge}(w,b)$ קמורה משאלה 3.a ולכן
$\frac{1}{m}\Sigma_{i=1}^m l_{x_i,y_i}^{hinge}(w,b)=\Sigma_{i=1}^m\frac{1}{m}l_{x_i,y_i}^{hinge}(w,b)$ קמורה משאלה 4.a.
$\frac{\lambda}{2}||w||^2$ קמורה מכיוון שהיא תבנית ריבועית.
לכן סכום f קמורה כסכום פונצקיות קמורות.

ii. נמצא איבר ב-subgradient שלה:

מלינאריות הנגזרת לפי w:

$$\Delta f(w,b) = \left(\frac{1}{m}\Sigma_{i=1}^{m} l_{x_i,y_i}^{hinge}(w,b)\right)' + \left(\frac{\lambda}{2}\left||w|\right|^2\right)'$$

$$= \frac{1}{m}\Sigma_{i=1}^{m}(l_{x_i,y_i}^{hinge}(w,b))' + \lambda w$$

$$= \lambda w - \Sigma_{i=1}^{m}\mathbf{1}_{[y_i\langle x_i,w\rangle<1]}x_i y_i$$