Recap.

Thm 3.8 of R & W (2018)

If $f_n \xrightarrow{e} f$. $\varepsilon_n \downarrow 0$.

$\limsup_n \varepsilon_n-\operatorname{argmin} f_n \subseteq \operatorname{argmin} f$.

$f_n \xrightarrow{e} f \iff \text{epi-}f_n \xrightarrow{PK} \text{epi-}f$

$\text{epi-}f = \{(x, \alpha) \in \mathcal{X} \times \mathbb{R} : f(x) \leq \alpha\}$.

We will define 'epi-distance' btw fcs by 'set-distance' btw epi-graphs.

Def. ρ-epidistance. $\forall \rho > 0$ ρ-AW distance.

$$dl_\rho(C, D) = \sup_{x \in B(0,\rho)} |d(x, C) - d(x, D)|$$

$$dl_\rho(f, g) = dl_\rho(\text{epi-}f, \text{epi-}g)$$

Rmk. dl_ρ is a pseudo metric: $dl_\rho(C, D) = 0 \not\Rightarrow C = D$.

dl_ρ is non-decreasing in ρ.

Def. Epi-distance.

$$dl(C, D) = \int_0^\infty e^{-\rho} dl_\rho(C, D) \, d\rho.$$

$$dl(f, g) = dl(\text{epi-}f, \text{epi-}g)$$

Thm 4.36 of R&W (1998)

$$C_n \xrightarrow{PK} C \iff dl_p(C_n, C) \to 0 \quad \forall p > 0.$$

$$\iff dl(C_n, C) \to 0.$$

Thm 7.58 of R&W (1998)

$$f_n \xrightarrow{e} f \iff dl(f_n, f) \to 0.$$

Moreover, $(\text{lsc-ftn}(\mathbb{R}^d), dl)$ is a proper, complete, separable metric space.

Rmk1. On reference point 0.

$\theta_0 = \arg\min f(\theta)$ $\Rightarrow$ $0 = \arg\min_u \left[f\left(\frac{u}{r_n} + \theta_0\right) - f(\theta_0)\right]$

$\hat{\theta}_n \in \arg\min f_n(\theta)$. $\hat{\theta}_n - \theta_0 = \arg\min_u \left[f_n\left(\frac{u}{r_n} + \theta_0\right) - f_n(\theta_0)\right]$

If we change the reference point, induced topology on $\text{lsc-ftn}(\mathbb{R}^d)$ doesn't change.

Rmk 2. (Pompeiu-Hausdorff distance)

$$dl_\infty(C, D) = \sup_x |d(x, C) - d(x, D)|$$

$dl_p \leq dl_\infty \quad \forall p > 0$. $\quad dl \leq dl_\infty$.

Therefore,

$$dl_\infty(C_n, C) \to 0 \Rightarrow C_n \xrightarrow{PK} C$$

In fact, convergence in dl_∞ is strictly stronger notion than PK-convergence.

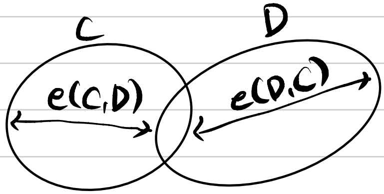
$a_n \to a$ $\quad |b_n| \to \infty$. $\quad C_n = \{a_n, b_n\}$ $\quad$ PK-$\lim_n C_n = \{a\}$.

$dl_p(C_n, C) \to 0$ $\quad dl_\infty(C_n, C) \to \infty$. If $C_n, C \subseteq B(0, R)$ then. ...

Def. Pseudo ρ-epi distance.

Def. ρ-truncated Hausdorff distance.

$$e(C, D) = \sup_{x \in C} d(x, D) = \sup_{x \in C} \inf_{y \in D} d(x, y)$$



$$d_{Haus}(C, D) = \max\{e(C,D), e(D,C)\}$$

$$\hat{dl}_\rho(C, D) = \max\{e(C \cap B(0,\rho), D),$$
$$e(D \cap B(0,\rho), C)\}$$

Rmk. $\hat{dl}_\rho$ is non-decreasing in ρ.

$\hat{dl}_\rho(C,D)$ is not a pseudo-metric. Triangle ineq does not hold.

Thm 4.36. $C_n \xrightarrow{PK} C \iff \hat{dl}_\rho(C_n, C) \to 0 \quad \forall \rho > 0$.

Prop 4.37 of R&W(1998) If C, D are closed,

$$\hat{dl}_\rho(C,D) \leq dl_\rho(C,D) \leq \hat{dl}_{\rho'}(C,D) \qquad \rho' = 2\rho + d(0,C) \vee d(0,D)$$

If C, D are convex, $\rho' = \rho + d(0,C) \vee d(0,D)$

If, C, D are convex, $0 \in C \cap D$. then $dl_\rho = \hat{dl}_\rho$

Thm 3.1 of Attouch and Wets (1993), Thm 4.1 of Royset (2020)

For any l.s.c. ftns f, g. Suppose that

$\inf f \in (-\rho, \rho)$, ε-$\operatorname{argmin} f \cap B(0, \rho) \neq \emptyset$.

Then,

$$\inf g - \inf f \leq \hat{dl}_\rho (f, g).$$

To be able to consider any perturbation of f and still obtain a proximity of minimizers, we need to know some geometrical shape at minimizer.

Thm 4.2 of Royset (2020). For any l.s.c. ftns f, g.

Suppose that $\inf f$, $\inf g \in (-\rho, \rho)$ and

$\operatorname{argmin} f \cap B(0, \rho) \neq \emptyset$, $\operatorname{argmin} g \cap B(0, \rho) \neq \emptyset$.

$$f(x) - \inf f \geq \varphi(\operatorname{dist}(x, \operatorname{argmin} f))$$

for some increasing $\varphi : [0, \infty) \to [0, \infty)$ s.t $\varphi(0) = 0$

Then,

$$e(\operatorname{argmin} g \cap B(0, \rho), \operatorname{argmin} f) \leq \hat{dl}_\rho (f, g) + \varphi^{-1}(\hat{dl}_\rho(f, g))$$

(Convexity) $f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle$

(Strong Convexity) $\exists \mu > 0$ s.t. $f(y) \geq f(x) + \langle \nabla f(x), y-x \rangle$
S.C.
$+ \frac{\mu}{2} \|y-x\|_2^2$

(Weak SC). $\exists \mu > 0$. $f(x_p) \geq f(x) + \langle \nabla f(x), x_p - x \rangle$
$+ \frac{\mu}{2} \|x_p - x\|_2^2$

$x_p = \mathrm{Proj}_{\mathrm{argmin} f}(x)$.

(Polyak-Łojasiewicz, PL.)

$\frac{1}{2} \|\nabla f(x)\|_2^2 \geq \mu (f(x) - \inf f)$

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

(Quadratic growth)

$f(x) - f^* \geq \frac{\mu}{2} \|x - x_p\|_2^2$

Thm 2 of Karimi et al. (2020)

(SC) $\to$ (WSC) $\to$ (PL) $\to$ (QG)

If f is diff'l, ∇f is Lipschitz, f is convex,

(QG) $\to$ (PL)

Ex. Many (linear) moment inequalities

θ^* solves

$$\mathbb{E}\, h_1(X,\theta) \le 0 \quad \vdots \quad \mathbb{E}\, h_J(X,\theta) \le 0$$

$$\boxed{\begin{aligned} -\mathbb{P}(X \ge \theta) + \tfrac{1}{2} &\le 0 \\ -\mathbb{P}(X \le \theta) + \tfrac{1}{2} &\le 0 \end{aligned}}$$

$$h_j(x,\theta) = a_j^T \theta - h_j(x)$$

$$f(\theta) = \max_{j=1..J} \left(a_j^T \theta - \mathbb{E}\, h_j(x) \right)_+$$

$$f_n(\theta) = \max_{j=1..J} \left(a_j^T \theta - \overline{h}_{j,n} \right)_+$$

$$\mathcal{A} = \{ \theta : A\theta \le h \text{ coordinate wise} \} \qquad A = \begin{bmatrix} a_1^T \\ \vdots \\ a_J^T \end{bmatrix} \qquad h = \begin{bmatrix} \mathbb{E}h_1(x) \\ \vdots \\ \mathbb{E}h_J(x) \end{bmatrix}$$

(Hoffman's constant)

$\exists\, K(A) > 0$

$\forall u$,

$$\mathrm{dist}_{\|\cdot\|_p}(u, \mathcal{A}) \lesssim K(A)_{p,q} \, \| (Au - h)_+ \|_q$$

$$f(u) - f^* \gtrsim K(A)_{p,\infty}^{-1} \, \mathrm{dist}_{\|\cdot\|_p}(u, \mathcal{A}) \quad \text{(linear conditioning)}$$

Thm 4.4 of Royset (2020).

For l.s.c fcns f_n, f. Suppose $f_n \xrightarrow{e} f$.

$\inf f_n \in (-\rho, \rho-\varepsilon)$, $\inf f \in (-\rho, \rho)$.

γ- argmin $f \cap B(0,\rho) \neq \emptyset$ $\forall \gamma > 0$.

Then,

$$e(\varepsilon\text{-}\operatorname{argmin} f_n \cap B(0,\rho), \delta\text{-}\operatorname{argmin} f) \leq \hat{dl}_\rho(f_n, f)$$

$\forall \delta > \varepsilon$

Prop 4.2 of Royset (2018, SIAM) For any l.s.c fcts.

$f, g : \mathcal{X} \to \overline{\mathbb{R}}$. $\forall \rho > 0$.

$$\hat{dl}_\rho(f,g) \leq \sup_{A_\rho} |f-g| \qquad A_\rho = \{x \in B(0,\rho) : f(x) \leq \rho \text{ or } g(x) \leq \rho\}$$

$$\leq \sup_{B(0,\rho)} |f-g|$$

Ex. Many linear moments ineq.

$$\hat{dl}_\rho(f, f_n) \leq \max_{j=1..J} |E h_j(x) - h_{j,n}(x)| \asymp \sqrt{\frac{\log J}{n}}$$

if $h_j(x)$'s are sub-g.