Mathematics in DL: **Feb 6 2018**

Lecture 4: Optimization

Lecturer: Juhee Lee *Scribes: Juhee Lee*

4.1 Optimization problem

최적화 문제를 정의하고 자주 쓰이는 미분을 이용한 최적화방법을 살펴봅니다.

Given a function f,

- 두 집합 X, R와 각각의 원소 $x_0 \in X, y \in R$에 대해 한 원소 x_0에 한 원소 $f(x_0)$가 대응될때, f를 X에서 R로의 실수함수라고 합니다.

$$\begin{aligned} f: \quad & X && \longrightarrow Y \\ & x && \longmapsto f(x) \end{aligned}$$

- 주어진 집합 X, R에 대해서 목적에 맞는 x_0를 찾는것이 목표입니다.
- $x_0 \in X$ such that $f(x_0) \leq f(x)$ for all $x \in X\}$ is called global "minimization"

  이때, objective function f를 "loss function" or "cost function"이라 합니다.
- $x_0 \in X$ such that $f(x_0) \geq f(x)$ for all $x \in X\}$ is called global "maximization"

  이때, objective function f를 "utility function" or "fitness function"이라 합니다
- 이때, x_0를 최적화 해 (optimal solution or feasible solution) 라고 합니다. †

Definition 4.1 *A multi-objective optimization problem is an optimization problem that involves multiple objective functions, $k \geq 2$*

$$\begin{aligned} f: \quad & X && \longrightarrow R^k \\ & x && \longmapsto (f_1(x), \cdots, f_k(x))^T \end{aligned}$$

find $x_0 \in X$ *such that* $\min(f_1(x), \cdots, f_k(x))$

Definition 4.2 *For above function f, x_0 is a"local minumum" if*

for all $x \in X$,

$$\exists \delta > 0, \quad ||x - x_0|| \Rightarrow f(x_0) \leq f(x)$$

Definition 4.3 *An extended real-valued function f is said to be "(Lebesgue)measuable " if its domain is measuable and if it satifies one of the following below.*

1. *For each real number α the set $\{x : f(x) > \alpha\}$ is measurable*

2. *For each real number* α *the set* $\{x : f(x) \geq \alpha\}$ *is measurable*
3. *For each real number* α *the set* $\{x : f(x) < \alpha\}$ *is measurable*
4. *For each real number* α *the set* $\{x : f(x) \leq \alpha\}$ *is measurable*
5. *For each real number* α *the set* $\{x : f(x) = \alpha\}$ *is measurable*

Note. The outer measure of an interval is its length
Note. For $A \subset I_n$,

$$m^* A = inf_{A \subset U I_n} \sum l(I_n)$$

Note. Given $c \in \mathbb{R}$, for $\epsilon > 0$, $N(c, \epsilon) = (c - \epsilon, c + \epsilon)$

Definition 4.4 *For* $O \subset \mathbb{R}$, O *is a"open set" if for all* $x \in O, \exists \epsilon > 0,\ N(x, \epsilon) \subset O$

Note. F is a"close set" if F^c is open set

Quiz. $(1, 3)$ is open set ? yes!

Quiz. $(1, 3]$ is open set?

Note. Interval $=\{(a, b), (a, \infty), [a, \infty), (-\infty, b], [a, b], \ldots)\}$

convex

Definition 4.5 *f is convex if* $\forall x_1,\ x_2 \in X, \forall t \in [0, 1]$,

$$f(tx_1 + (1 - t)x_2 \leq tf(x_1) + (1 - t)f(x_2) : \textit{Jensen's inequality}$$

If X is random variable of f then $f(E(X)) \leq E(f(x))$

Definition 4.6 *f is concave if* $\forall x_1,\ x_2 \in X, \forall t \in [0, 1]$,

$$f(tx_1 + (1 - t)x_2 \geq tf(x_1) + (1 - t)f(x_2)$$

Note. Any local minimum of a convex function is also a global minimum.

Note. A practical test for convexity if its second derivative is non-negative ex. Check $f(x) = x^2$ is convex sol) $f''(x) = 2 > 0$ so f is convex. Note. If $f(x), g(x)$ are convex functions

1. $m(x) = \max(f(x), g(x)), h(x) = [f(x) + g(x)]$ convex functions
2. $f(x)$ is concave, $g(x)$ is non-decreasing then $f(f(x))$ is convex function

Affine space

Note. $\lambda\mathbf{a} + (1 - \lambda)\mathbf{b}$,
Note. $\mathbf{p} + \lambda(\mathbf{a} - \mathbf{p}) + (1 - \lambda)(\mathbf{p} - \mathbf{b})$,

Review of Derivatives

Definition 4.7 *A function f is differentiable at a if $f'(a)$ exists.*

Note. If f is differentiable at a, then f is continuous at a

Theorem 4.8 *Basic differentiation formulas*

1. $\frac{d}{dx}(c) = 0$
2. $\frac{d}{dx}(cf) = c\frac{df}{dx}$
3. $\frac{d}{dx}(f \pm g) = \frac{df}{dx} \pm \frac{dg}{dx}$ *: linearity*
4. $\frac{d}{dx}(fg) = f\frac{dg}{dx} + g\frac{df}{dx}$ *: product rule*
5. $\frac{d}{dx}f(g(x)) = \frac{df(g(x))}{dx} \cdot \frac{dg}{dx}$ *: chain rule*
6. $\frac{d}{dx}(x^n) = nx^{n-1}$ *: power rule (n is any real number)*
7. $\frac{d}{dx}(\sin x) = \cos x$
8. $\frac{d}{dx}(\cos x) = -\sin x$
9. $\frac{d}{dx}(a^x) = a^x lna$
10. $\frac{d}{dx}(lnx) = \frac{1}{x}$

4.2 First order optimization

ex. $f(x) = x + 2,\ x \in [1, 2]$ 에 대하여, 최대점, 최소점을 구하여라. sol) $x = 1, f(x) = 3$ 즉 $(1, 3)$ 최소점. $x = 2, f(x) = 4$ 즉 $(2, 4)$ 최대점.

Theorem 4.9 *Weierstrass theorem*
F is continuous in $[a, b]$, f is bounded in $[a, b]$ and has a maximum and minimum.

Theorem 4.10 *Fermat theorem*
f has local maximum or local minimum at (a, b), there exists $f'(c)$, $f'(c) = 0$.

ex. $f(x) = x^3 - 1$, find a local minimum or maximum. sol) no

Definition 4.11 *critical point*
c is a critical point in X such that $f'(c) = 0$ or $f'(c)$ does not exists

ex. $f(x) = x^{2/3}$, find a critical point.
sol) $x = 0$

Theorem 4.12 *mean value theorem*
f is continuous in $[a, b]$, f is diffentiable in (a, b), there exists c such that $f'(c) = \frac{f(b)-f(a)}{b-a}$

Definition 4.13 *inflection point*
a point x on curve such that $f''(x - \epsilon)f''(x + \epsilon) < 0$

4.3 Second order optimization

Using Newton Method,

Theorem 4.14 *newton method*
$y = f(x), (a_1, f(a_1),$
$Using\ y - f(a_1) = f'(a_1)(x - a_1)$
$a_{n+1} = a_n - \frac{f(a_n)}{f'(a_n)}$

References

[1] Essential Calculus, Early Transcendentals, Janmes Stewart

[2] https://en.wikipedia.org

[3] $https\ :\ //www.khanacademy.org/math/ap - calculus - bc/bc - applications - derivatives_bc - optimization$