Slides for the MFAI (Aug-Dec 2024) Lectures Sep 25 (11am-12pm) , Sep 28 (14:40-16:40)

C R Subramanian

CSE Dept, SECS, Indian Institute of Technology, Bhubaneswar.

Example Motivation :

- Given : $P = \{(\vec{x}_i, y_i) : \vec{x}_i \in \mathcal{R}^d, y_i \in \mathcal{R})\}_{i=1,\ldots,n}$;
- Find : a function $y = f(\vec{x}) = \vec{a} \cdot \vec{x}$, $\vec{a} \in \mathcal{R}^d$, minimising
- total sum of squares of errors $E(\vec{a}) = \sum_{i=1}^{n} (y_i - \vec{a} \cdot \vec{x}_i)^2$.
-
- Want to find a $\vec{a} \in \mathcal{R}^d$ which minimises $E(\vec{a})$.
-
- When $d = 1$, $E(a)$ becomes a continuous function of one variable a.
- The minimiser a^* and the minimum value $E(a)$ can be computed in $O(n)$ time.

Limits and Continuity

- $f : O \to \mathcal{R}$ is a function. O is an open set. Let $a \in O$.
- limit of $f(x)$ as x approaches a is L if
- $\forall \epsilon > 0 \; \exists \delta > 0$ such that $0 < |x - a| < \delta \Rightarrow |f(x) - L| \leq \epsilon$.
- Denoted by : $Lt_{x \to a} \, f(x) = L$.
- Left limit : $Lt_{x \to a^-} f(x) = L$. $(-\delta < x - a < 0)$
- Right limit : $Lt_{x \to a^+} f(x) = L$. $(0 < x - a < \delta)$.
-
- L exists if and only if left- and right- limits exist and equal L.
- Example : $f(x) = [x]$ does not have limits when x is an integer ; both left- and right- limits of f exist at integers.
- $f(x) = 1/x$ has limits everywhere but not at $x = 0$. Both left and right limits do not exist at $x = 0$.

Limits and Continuity

- f is *continuous* at a if $f(a)$ is defined and $Lt_{x \to a} f(x) = f(a)$.
- $x, x^2, x^3, \sin(x), \cos(x), e^x, |x|$ - continuous everywhere.
- $f(x) = [x]$ continuous everywhere except at integers
- $f(x) = x^{-1}$ continuous everywhere except at $x = 0$.
- f and g are continuous at a. Then, $f + g$, $f - g$, $f \cdot g$ are continuous at a. $g(a) \neq 0 \Rightarrow f/g$ cont. at a.
-
- f is continuous at a, $\quad g$ is continuous at $f(a) \quad \Rightarrow$ $h(x) = g(f(x))$ is continuous at a.
- $sin\left(e^{x^2}\right)$, $e^{sin(x^2)}$ and $\left(e^{sin(x)}\right)^2$ are continuous everywhere.
-
- f is cont. over $[a, b]$ with $f(a) < f(b)$. Then, $\forall c \in (f(a), f(b)) \;\; \exists x \in (a, b)$ such that $f(x) = c$.
- f is continuous over $[a, b]$ implies f is bounded over $[a, b]$.
- f is continuous over $[a, b]$ implies f achieves its min and max.

Differentiability

- f is *differentiable* at a if $Lt_{x \to a} \frac{f(x)-f(a)}{x-a}$ exists.
- limit is the *derivative* of f at a, denoted by
- $f'(a)$, $f^{(1)}(a)$, $\frac{df(a)}{dx}$.
- $x, x^2, x^3, e^x, \sin(x), \cos(x)$ -differentiable at every $x \in \mathcal{R}$.
- $|x|$ is differentiable everywhere except at $x = 0$.
-
- f is differentiable at $a \Rightarrow f$ is continuous at a.
- Converse need not be true : $|x|$ and $x = 0$, for example.
- Left-derivative : same except we focus on $x < a$.
- Right-derivative : same except we focus on $x > a$.
- For $|x|$, $f_L'(0) = -1$ and $f_R'(0) = +1$.

Differentiability

- **Algebra :**
- f and g are defined over $\mathcal{R}$.
- $f'(a)$ and $g'(a)$ exist for $a \in \mathcal{R}$.
- $(f \pm g)'(a) = f'(a) \pm g'(a)$.
- $(f \cdot g)'(a) = f(a) \cdot g'(a) + f'(a) \cdot g(a)$.
- $\left(\frac{f}{g}\right)'(a) = \frac{g(a) \cdot f'(a) - f(a) \cdot g'(a)}{g(a)^2}$ provided $g(a) \neq 0$.
-
- **Chain Rule :**
- Suppose $Range(f) \subseteq Domain(g)$; $f'(a)$, $g'(f(a))$ exist.
- $(g(f))'(a)$ exists and equals $g'(f(a)) \cdot f'(a)$.
- Familiar version :
- $y = f(x)$, $z = g(y)$, $z = g(f(x)) \Rightarrow \frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$.

Differentiability

- For $x \in \mathcal{R}$, $B(x, \delta) := \{y \in \mathcal{R} : 0 \leq |y - x| < \delta\}$.
- f is *twice-differentiable* at a if
- (i) for some $\delta > 0$, $f'(x)$ exists for every $x \in B(a, \delta)$
- (ii) derivative of $f'(x)$ $(= Lt_{x \to a} \frac{f'(x) - f'(a)}{x - a})$ exists at a.
- Second derivative is denoted by $f''(a)$, $f^{(2)}(a)$ and $\frac{df^2(a)}{dx^2}$.
-
- Generally, for $k \geq 1$, f is *k-times differentiable* at a if
- (i) for some $\delta > 0$, $f^{(k-1)}(x)$ exists for every $x \in B(a, \delta)$
- (ii) $f^{(k-1)}(x)$ is differentiable at a.
- k-th derivative denoted by $f^{(k)}(a)$ or $\frac{df^k(a)}{dx^k}$.

Differentiability

- $x, x^2, x^3, e^x, \sin(x), \cos(x)$ - k-times differentiable for every $k \geq 1$ and everywhere.
- $f(x) = \log_e x$ - $f^{(k)}(x)$ exists for every $k \geq 1$ for every $x > 0$.
-
- a is a local minimum / local maximum of f if
- $f(a) \leq f(x)$ / $f(a) \geq f(x)$
- for every $x \in B(a, \delta)$ for some $\delta > 0$.
-
- $f : O \rightarrow \mathcal{R}$, O is open.
- $a \in O$ is a global minimum / global maximum of f over O if
- $f(a) \leq f(x)$ / $f(a) \geq f(x)$ for every $x \in O$.
- Every global optimum is also a local optimum.

Differentiability and optima

- If a is a local optimum for f, then $f'(a) = 0$.
-
- Necessary but not sufficient.
- Example : $f(x) = x^3$ for $x < 0$ and $f(x) = x^2$ for $x \geq 0$.
- $f'(0) = 0$ but 0 is neither a local minimum nor a local maximum for f.
-
- a is a *saddle point* if $f'(a) = 0$ but a is not a local optimum.
- $f'(a) = 0$ - a is a critical point.

Differentiability and optima

- $f'(a) = 0$ and $f''(a) > 0 \Rightarrow a$ is a local minimum for f.
- sufficient but not necessary.
- Eg : $f(x) = -x^3$ for $x \leq 0$ and $f(x) = x^3$ for $x > 0$.
- 0 is global minimum for f. But, $f'(0) = f''(0) = 0$.
-
- $g'(a) = 0$ and $g''(a) < 0 \Rightarrow a$ is a local maximum for g.
- sufficient but not necessary.
- Eg : $g(x) = -f(x)$
- 0 is global maximum for g. But, $g'(0) = g''(0) = 0$.
-

Taylor's Approximation Formula

- f'' exists and is continuous over $(a-\delta, a+\delta)$ for some $\delta > 0$.
- **Taylor's first-order approximation formula :**
- $f(x) = f(a) + f'(a)(x-a) + E_1(x)$, $\forall x \in B(a, \delta)$
- where $E_1(x) = \int_a^x (x-t)f''(t)dt \to 0$ as $x \to a$.
-
- $E_1(x) = \frac{f''(c)(x-a)^2}{2}$ for some $c \in (a, x)$.
- $f(a+h) = f(a) + hf'(a) + o(h)$ as $h \to 0$.
- $f(a+h) \approx f(a) + hf'(a)$ as $h \to 0$.
-
- differentiability $\iff$ local linearizability.

Taylor's Approximation Formula

- f''' exists and is continuous over $(a-\delta, a+\delta)$ for some $\delta > 0$.
- **Taylor's second-order approximation formula :**
- $f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)(x-a)^2}{2} + E_2(x), \forall x \in B(a, \delta)$
- where $E_2(x) = \frac{1}{2} \cdot \int_a^x (x-t)^2 f'''(t)dt \to 0$ as $x \to a$.
-
- $E_2(x) = \frac{f'''(c)(x-a)^3}{6}$ for some $c \in (a, x)$.
- $f(a+h) = f(a) + hf'(a) + \frac{h^2 f''(a)}{2} + o(h^2)$ as $h \to 0$.
- $f(a+h) \approx f(a) + hf'(a) + \frac{h^2 f''(a)}{2}$ as $h \to 0$.

Taylor's Approximation Formula

- $f^{(n+1)}()$ exists, continuous over $(a-\delta, a+\delta)$ for some $\delta > 0$.
- **Taylor's nth-order approximation formula :**
- $f(x) = \sum_{j=0}^{n} \frac{f^{(j)}(a)(x-a)^j}{j!} + E_n(x)$, $\forall x \in B(a, \delta)$
- where $E_n(x) = \frac{1}{n!} \cdot \int_a^x (x-t)^n f^{(n+1)}(t)dt \to 0$ as $x \to a$.
- $f^{(0)}(a) = f(a)$.
-
- $E_n(x) = \frac{f^{(n+1)}(c)(x-a)^{n+1}}{(n+1)!}$ for some $c \in (a, x)$.
- $f(a+h) = \sum_{j=0}^{n} \frac{f^{(j)}(a)h^j}{j!} + o(h^n)$ as $h \to 0$.
- $f(a+h) \approx \sum_{j=0}^{n} \frac{f^{(j)}(a)h^j}{j!}$ as $h \to 0$.

Taylor's series

- f is infinitely differentiable over $(a-\delta, a+\delta)$ for some $\delta > 0$.
- **Taylor series expansion for $f(x)$:**
- $f(x) = f(a) + f'(a)(x-a) + \ldots + \frac{f^{(n)}(a)(x-a)^n}{n!} + \ldots$.
-
- $f(x) = \sum_{j=0}^{\infty} \frac{f^{(j)}(a)(x-a)^j}{j!}, \forall x \in B(a, \delta)$
-
- $f(a+h) = \sum_{j=0}^{\infty} \frac{f^{(j)}(a)h^j}{j!}, \forall h \in (-\delta, \delta)$.
-