Class #14

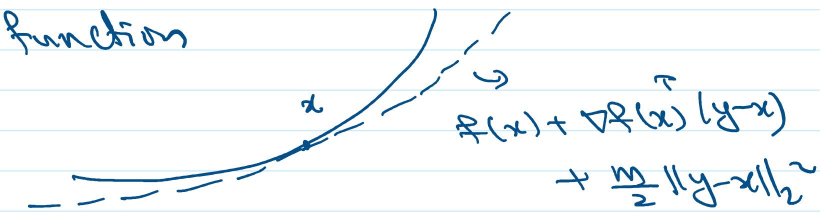
Friday, March 7, 2025 10:19 AM

③ A function f is strongly convex with parameter $m > 0$ if and only if

$$\Leftrightarrow \quad f(y) \geq f(x) + \nabla f(x)^T (y-x) + \frac{m}{2}\|y-x\|_2^2$$

$$\forall\, x, y \in \operatorname{dom} f$$

i.e., we have a quadratic function that globally underestimates our function



④ A function f is strongly convex with parameter $m > 0$ if and only if

$$g(x) = f(x) - \frac{m}{2}\underbrace{\|x\|_2^2}_{x^T x} \quad \text{is convex.}$$

Let $h(x)$ be a convex function

$\Rightarrow$ $h(x) + \frac{m}{2}\|x\|_2^2$ is strongly convex with parameter m.

⑤ Strong monotonicity of gradient

f is strongly convex with parameter m iff

$$(\nabla f(x) - \nabla f(y))^T (x-y) \geq m\|x-y\|^2$$

$$\left(\nabla f(x) - \nabla f(y)\right)^T (x-y) \geq m\|x-y\|_2^2$$
$\forall\, x, y \in \operatorname{dom} f$

Basic Real Analysis (continued)

Closed Set

Let C be a set in $\mathbb{R}^n$. We say C is closed iff $\mathbb{R}^n \setminus C = \{x \in \mathbb{R}^n : x \notin C\}$ is open.

① $C = (-\infty, 1) \Rightarrow \mathbb{R}\setminus C = \underbrace{[1, \infty)}_{\text{not open}}$

$\Rightarrow C$ is not closed.

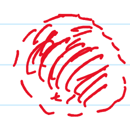
② $C = (-\infty, 1] \Rightarrow \mathbb{R}\setminus C = \underbrace{(1, \infty)}_{\text{open}}$

$\Rightarrow C$ is closed.

③ $[1, 2] \Rightarrow$ closed

$[1,3] \cup [5,6] \Rightarrow$ closed

closed



open

Suppose $x_1, x_2, \ldots$ is a sequence in set C and let $x_n \xrightarrow{n\to\infty} x^*$.

$\hookrightarrow$ limit point of the sequence

C is a closed set iff every limit point belongs to C (i.e., $x^* \in C$)

e.g.; $C = (0,1)$; $C = (0,1] \Rightarrow$ neither open nor closed

$x_n = \frac{1}{n} \to 0 \notin C$

$x_n = 1 - \frac{1}{n} \to 1 \notin C \Rightarrow C$ is not closed

$C = [0,1] \Rightarrow$ closed set.

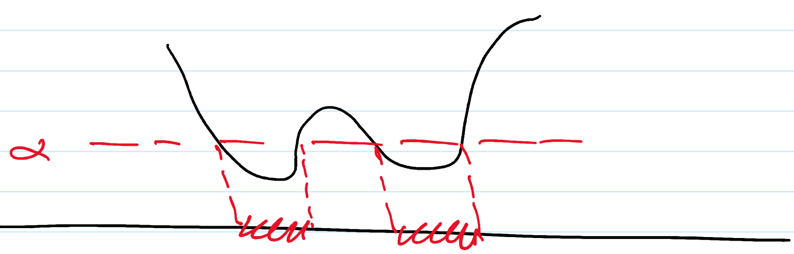
* ϕ is both open and closed
* $\mathbb{R}^n$ is both open and closed

Sublevel set of a function

Let $f : \mathbb{R}^n \to \mathbb{R}$. Let $\alpha \in \mathbb{R}$.

The $S_\alpha = \{x \in \operatorname{dom} f : f(x) \le \alpha\}$

is called α-sublevel set of f.



In optimization, when using descent methods in particular, we are interested in the sublevel set produced by the initialization.

$$S = \{x \in \text{dom} f : f(x) \leq f(x^{(0)})\}$$

where $x^{(0)}$ is my initialization.

We need that the initial sublevel set S is closed for the methods to generate a convergent sequence.

Functions whose every sublevel set is closed are called called closed functions.

Condition for closedness of functions

① If $f : \mathbb{R}^n \to \mathbb{R}$ is continuous and dom f is closed $\Rightarrow$ f is closed.

② If $f : \mathbb{R}^n \to \mathbb{R}$ is continuous but dom f is open then f is closed if and only if

whenever we have a sequence $\{x_i \in \text{dom} f\}$ that converges to a point on the boundary of the dom f

i.e. $\lim_{i \to \infty} x_i = x \in \underbrace{\text{bd}}_{\text{boundary}} \text{dom} f$

then $\lim_{i \to \infty} f(x_i) = \pm\infty$

e.g.; $f(x) = x \log x$; dom $f = (0, \infty)$

e.g.; $f(x) = x\log x$; $\text{dom} f = (0,\infty)$

↳ Not a closed function

$$x_i = \frac{1}{i} \to 0$$

but $\lim_{i\to\infty} f(x_i) = 0$

e.g.; $f(x) = -\log x$, $\text{dom} f = (0,\infty)$

$$x_i = \frac{1}{i} \to 0$$

$\lim_{i\to\infty} f(x_i) = \infty \Rightarrow f(x)$ is closed.

Going forward, we will assume that

$$f \in C^2(\mathbb{R}^n)$$

↳ twice continuously differentiable

and $S = \{x \in \text{dom} f : f(x) \le f(x^{(0)})\}$ is closed. [If f is closed, this is obvious]

In the case of strongly convex functions, it turns out that these conditions are enough to imply that:

* $f(x)$ has Lipschitz continuous gradients on the sublevel set S (locally Lipschitz continuity of gradients)

$\Uparrow$

gradients)

$\Updownarrow$

★ $f(x)$ is also upper bounded by a quadratic function on the set S.

we will show this by establishing that the maximum eigenvalue of $\nabla^2 f(x)$ within the set S is upper bounded, which we denote by M.

$$mI \preceq \nabla^2 f(x) \preceq MI \quad \forall\, x \in S.$$

[Note: Since $\nabla^2 f(x) \preceq MI \iff \lambda_{max}(\nabla^2 f(x)) \leq M \;\; \forall x \in S$

$$\Rightarrow \|\nabla f(x) - \nabla f(y)\|_2 \leq M \|x-y\|_2 \quad \forall\, x,y \in S$$

i.e. M is the Lipschitz continuity parameter.]

Lemma 1: Since f is s.c.

$\Rightarrow$ Set S is bounded

i.e. $\forall\, y \in S$, $\|y\|_2 \leq B$ for some $B \in \mathbb{R}$.

Proof:

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) + \frac{m}{2}\|y-x\|_2^2$$

Take $x = x^* \rightarrow x^* = \arg\min_{x \in \mathrm{dom} f} f(x)$

Clearly $x^* \in S$.

$\nabla f(x^*) = 0$

$$\Rightarrow \quad f(y) \geq f(x^*) + \frac{m}{2}\|y - x^*\|_2^2$$

$$\forall\, y \in S, \quad f(y) \leq f(x^{(0)})$$

$$\Rightarrow \quad f(x^{(0)}) \geq f(x^*) + \frac{m}{2}\|y - x^*\|_2^2 \quad \forall\, y \in S.$$

$$\underbrace{\sqrt{\frac{2\left(f(x^{(0)}) - f(x^*)\right)}{m}}}_{\text{constant} \,=\, B'} \geq \|y - x^*\|_2$$

$$B' \geq \|y - x^*\|_2 \geq \|y\|_2 - \|x^*\|_2$$

$$\Leftrightarrow \|y\|_2 \leq \underbrace{B' + \|x^*\|_2}_{B}$$

$$\Leftrightarrow \quad \|y\|_2 \leq B \quad \forall\, y \in S$$

$\boxtimes$

Compact in $\mathbb{R}^n$

Thus, S is both closed and bounded for strongly convex functions.

Fact from real analysis: A continuous function on a compact set is always bounded.

Look at $f(x) = -\log x$ for $x \in (0,1]$

but $f(x) \to \infty$ as $x \to 0$

Let us consider $g(x) = \lambda_{max}(\nabla^2 f(x))$ on $x \in S$

$g(x) : \mathbb{R}^n \to \mathbb{R}$

S is compact (for f s.c.).

g is continuous (since $\lambda_{max}(\cdot)$ is a continuous fn., $\nabla^2 f(x)$ is continuous b/c $f \in C^2$)

$\Rightarrow g$ is bounded on $S \Rightarrow \lambda_{max}(\nabla^2 f(x)) \leq M$ $\forall x \in S$.

i.e. $\forall x \in S$, $mI \preceq \nabla^2 f(x) \preceq MI$

Condition number of a strongly convex function on S

$$K = \text{cond}(f) = \frac{M}{m}$$

$\hookrightarrow$ Kappa (condition number) dictates the difficulty of an optimization problem.

$\hookrightarrow$ Larger $K \Rightarrow$ More iterations needed

Smaller $K \Rightarrow$ Less iterations needed

Quadratic upper bound on strongly convex f in S

Taylor's theorem with remainder

$$f(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(z)(y-x)$$

$$f(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2} \underbrace{(y-x)^T \nabla^2 f(z) (y-x)}_{\leq M \|y-x\|_2^2 \;\; \forall x,y \in S}$$

for some $z \in [x, y]$

$$\Rightarrow f(y) \leq f(x) + \nabla f(x)^T (y-x) + \frac{M}{2} \|y-x\|_2^2$$

$\Rightarrow$ A strongly convex function is both lower and upper bounded by a quadratic on the set S. $\blacksquare$