1 Cheat Sheet

1.1 Lipschitz Continuous

Definition 1. *$f \in C_L^{k,p}(Q)$ is k times continuously differentiable on Q if for all $x, y \in Q$,*

$$\|\nabla^p f(y) - \nabla^p f(x)\| \leq L \|y - x\| \tag{1}$$

$f \in C_L^{1,1}(\mathbb{R}^n)$ is continously differentiable on $\mathbb{R}^n$ if for all $x, y \in \mathbb{R}^n$

$$\|\nabla f(y) - \nabla f(x)\| \leq L \|y - x\| \tag{2}$$

Definition 2. *$f \in C_L^{2,1}(\mathbb{R}^n) \subset C_L^{1,1}(\mathbb{R}^n)$ if for all $x \in \mathbb{R}^n$, either condition is satisfied*

$$\left\|\nabla^2 f(x)\right\| \leq L \tag{3}$$

$$-L\mathbf{I} \preceq \nabla^2 f(x) \preceq L\mathbf{I} \tag{4}$$

Property for $f \in C_L^{1,1}(\mathbb{R}^n)$,

$$|f(y) - f(x) - \langle \nabla f(x), y - x\rangle| \leq \frac{L}{2} \|y - x\|^2 \tag{5}$$

$$f(x - \alpha \nabla f(x)) \leq f(x) - \alpha\left(1 - \frac{\alpha}{2} L\right) \|\nabla f(x)\|^2 \tag{6}$$

$$f\left(x - \frac{1}{L}\nabla f(x)\right) \leq f(x) - \frac{1}{2L} \|\nabla f(x)\|^2 \tag{7}$$

Note (5) implies that quadratic functions $\phi_-(\cdot), \phi_+(\cdot)$ are global lower/upper bound of $f(\cdot)$ respectively, i.e. $\phi_-(y) \leq f(y) \leq \phi_+(y)$ for any $x \in \mathbb{R}^n$, where

$$\phi_-(y) = f(x) + \langle \nabla f(x), y - x\rangle - \frac{L}{2}\|y - x\|^2$$

$$\phi_+(y) = f(x) + \langle \nabla f(x), y - x\rangle + \frac{L}{2}\|y - x\|^2$$

Note (7) is important in proving convergence of descent methods, where we see progress made in reducing function value of iterates by steping in $-\frac{1}{L}\nabla f(x)$ is at least some constant times the gradient norm.
Property for $f \in C_M^{2,2}(\mathbb{R}^n)$,

$$\left\|\nabla f(y) - \nabla f(x) - \nabla^2 f(x)(y - x)\right\| \leq \frac{M}{2}\|y - x\|^2 \tag{8}$$

$$\nabla^2 f(x) - C \preceq \nabla^2 f(y) \preceq \nabla^2 f(x) + C \quad \text{where} \quad c = M\|y - x\|\mathbf{I} \tag{9}$$

1.2 Convex

Definition 3. *The following are equivalent*

1. *A continuously differentiable function f is convex on convex set Q ($f \in \mathscr{F}^1(Q)$)*
2. *For all $x, y \in Q$*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle \tag{10}$$

3. *For all $x, y \in Q$ and $\lambda \in [0,1]$,*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y) \tag{11}$$

4. *For all $x, y \in Q$,*

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle > 0 \tag{12}$$

Definition 4. *A twice differentiable function f belongs to $\mathscr{F}^2(Q)$ if for any $x \in Q$,*

$$\nabla^2 f(x) \succeq 0 \tag{13}$$

1.3 Smooth & Convex

Definition 5. *$f \in \mathscr{F}_L^{1,1}(Q, \|\cdot\|)$ if f is convex with Lipschitz continous gradient, i.e. for all $x, y \in Q$,*

$$\|\nabla f(x) - \nabla f(y)\|_* \leq L \|x - y\| \tag{14}$$

Property of $f \in \mathscr{F}^{1,1}(\mathbb{R}^n, \|\cdot\|)$. Let $x, y \in \mathbb{R}^n$ and $\lambda \in [0,1]$

$$0 \leq f(y) - f(x) - \langle \nabla f(x), y - x \rangle \leq \frac{L}{2} \|x - y\|^2 \tag{15}$$

$$f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2L} \|\nabla f(x) - \nabla f(y)\|_*^2 \leq f(y) \tag{16}$$

$$\frac{1}{L} \|\nabla f(x) - \nabla f(y)\|_*^2 \leq \langle \nabla f(x) - \nabla f(y), x - y \rangle \tag{17}$$

$$0 \leq \langle \nabla f(x) - \nabla f(y), x - y \rangle \leq L \|x - y\|^2 \tag{18}$$

Note (15) implies a tighter (linear) lower bound to $f(\cdot)$ if we assume convexity. In fact, the lower bound can be improved further to a upward quadratic by (16).

1.4 Strongly Convex

Definition 6. *A continuously differentiable function f is strongly convex on $\mathbb{R}^n$ ($f \in \mathscr{S}_\mu^1(Q, \|\cdot\|)$) if there exists a convexity parameter $\mu > 0$ such that for all $x, y \in Q$,*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2} \|y - x\|^2 \tag{19}$$

Property for $f \in \mathscr{S}_\mu^1(\mathbb{R}^n)$. Let $x, y \in Q$ and $\lambda \in [0,1]$,

$$f(x) \geq f(x^*) + \frac{\mu}{2} \|x - x^*\|^2 \quad \text{where} \quad \nabla f(x^*) = 0 \tag{20}$$

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \mu \|x - y\|^2 \tag{21}$$

$$\lambda f(x) + (1-\lambda) f(y) \geq f(\lambda x + (1-\lambda)y) + \lambda(1-\lambda)\frac{\mu}{2} \|x - y\|^2 \tag{22}$$

1.5 Smooth & Strongly Convex

Definition 7. *A continuously differentiable function f that is strongly convex with L-lipschitz continuous gradients ($f \in \mathscr{S}_{L,\mu}^{1,1}(\mathbb{R}^n)$). Note $\kappa = L/\mu \geq 1$ is the condition number of f.*

Property for $f \in \mathscr{S}_{\mu,L}^{1,1}(\mathbb{R}^n)$. For any $x, y \in \mathbb{R}^n$

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \frac{\mu L}{\mu + L} \|x - y\|^2 + \frac{1}{\mu + L} \|\nabla f(x) - \nabla f(y)\|^2 \tag{23}$$