凸函数的性质（在 $\mathbb{R}^n$ 上）

Convex + Lipschitz: $\mathcal{S}_{0,L}^1(\mathbb{R}^n)$ 函数的性质

Thm 2.1.6 （用$\nabla^2 f(x)$表示的$\mathcal{S}_{0,L}^1(\mathbb{R}^n)$的等价条件）

$$f \in \mathcal{S}_{0,L}^1(\mathbb{R}^n) \iff 0 \leq \nabla^2 f(x) \leq L I_n, \quad \forall x \in \mathbb{R}^n \tag{7}$$

> 证明同Thm 2.1.6$'$

Thm 2.1.5 若 $f \in \mathcal{S}_{0,L}^1(\mathbb{R}^n)$，则 $\forall x, y \in \mathbb{R}^n$ 有

$$\frac{1}{2L}\|\nabla f(y) - \nabla f(x)\|_*^2 \leq f(y) - f(x) - \langle \nabla f(x),\ y - x\rangle \leq \frac{L}{2}\|y - x\|^2 \tag{8.1}$$

$$\frac{1}{L}\|\nabla f(y) - \nabla f(x)\|_*^2 \leq \langle \nabla f(y) - \nabla f(x),\ y - x\rangle \leq L\|y - x\|^2 \tag{8.2}$$

$$\frac{\alpha(1-\alpha)}{2L}\|\nabla f(y) - \nabla f(x)\|_*^2 \leq \alpha f(x) + (1-\alpha) f(y) - f\left(\alpha x + (1-\alpha) y\right) \leq \frac{\alpha(1-\alpha)L}{2}\|y - x\|^2 \tag{8.3}$$

> 证明：见笔记<函数性质证明>

Strongly Convex: $\mathcal{S}_\mu(\mathbb{R}^n)$ 函数的性质

Cor 2.1.8 若 $f \in \mathcal{S}_\mu(\mathbb{R}^n)$ 且 $\nabla f(x^*) = 0$，则 $\forall x \in \mathbb{R}^n$ 有

$$f(x) \geq f(x^*) + \frac{1}{2}\mu\|x - x^*\|^2 \tag{9.0}$$

> 证明：由(2.2)，
> $f(x) \geq f(x^*) + \langle \nabla f(x^*),\ x - x^*\rangle + \frac{\mu}{2}\|x - x^*\|^2 = f(x^*) + \frac{1}{2}\mu\|x - x^*\|^2$ $\square$.

Thm 2.1.10 若 $f \in \mathcal{S}_\mu(\mathbb{R}^n)$，则 $\forall x, y \in \mathbb{R}^n$ 有

$$\mu\,\|x - y\| \leq \|\nabla f(x) - \nabla f(y)\| \tag{9.1}$$

$$\frac{\mu}{2}\|y - x\|^2 \leq f(y) - f(x) - \langle \nabla f(x),\ y - x\rangle \leq \frac{1}{2\mu}\|\nabla f(y) - \nabla f(x)\|_*^2 \tag{9.2}$$

$$\mu\|y - x\|^2 \leq \langle \nabla f(y) - \nabla f(x),\ y - x\rangle \leq \frac{1}{\mu}\|\nabla f(y) - \nabla f(x)\|_*^2 \tag{9.3}$$

$$\frac{\alpha(1-\alpha)\mu}{2}\|y - x\|^2 \leq \alpha f(x) + (1-\alpha) f(y) - f\left(\alpha x + (1-\alpha) y\right) \tag{9.4}$$

> 证明：见笔记<函数性质证明>
>
> 推广：类似Con+ Lip

Strongly Convex + Lipschitz: $\mathcal{S}^1_{\mu,L}(\mathbb{R}^n)$ 函数的性质

Thm 2.1.12 若 $f\in\mathcal{S}^1_{\mu,L}(\mathbb{R}^n)$，则 $\forall x,y\in\mathbb{R}^n$ 有

$$\langle\nabla f(y)-\nabla f(x),\ y-x\rangle\le\frac{\mu L}{\mu+L}\|y-x\|^2+\frac{1}{\mu+L}\|\nabla f(y)-\nabla f(x)\|^2\tag{10}$$

> 证明：注意到一定有 $\mu\le L$。如果 $\mu=L$，则 $\nabla^2 f(x)\equiv\mu=L$，此时 $f(x)=\frac{\mu}{2}\|x-x^*\|^2$，验证可知(10)式 $\text{左}=\mu\|y-x\|^2=\text{右}$。
>
> 如果 $\mu<L$，令 $h(x)=f(x)-\frac{\mu}{2}\|x\|^2\in\mathcal{S}^1_{0,L-\mu}(\mathbb{R}^n)$，由(8.2)可得
>
> $$\langle\nabla h(y)-\nabla h(x),\ y-x\rangle\ge\frac{1}{L-\mu}\|\nabla h(y)-\nabla h(x)\|^2$$
>
> 即
>
> $$\langle\nabla f(y)-\nabla f(x),\ y-x\rangle-\mu\|y-x\|^2\ge\frac{1}{L-\mu}\|\nabla f(y)-\nabla f(x)-\mu(y-x)\|^2$$
>
> 化简即可得(10) □.

Thm A.3.

$$f\in\mathcal{S}^1_{\mu,L}(\mathbb{R}^n)\iff$$
$$f(y)-f(x)-\langle\nabla f(x),\ y-x\rangle\ge\frac{1}{2L}\|\nabla f(y)-\nabla f(x)\|^2+\frac{\mu}{2(1-\mu/L)}\left\|y-x-\frac{1}{L}(\nabla f(y)-\nabla f(x))\right\|^2$$

> 证明略

容易看出，在无约束条件下多出的性质一般是与函数梯度差的范数相关的性质。其中凸Lipschitz函数的性质与强凸函数的性质在结果上具有很强的对称性。

参考文献

本文主要结论参考

1. *Lectures on Convex Optimization* by Yurii Nesterov
2. *Acceleration Methods* by Alexandre d’Aspremont