OP202 Homework

Yanyu ZHOU

January 31, 2023

1

Lemma 3: Function f belongs to $\mathcal{C}_L^{2,1}(\mathbf{R}^n)$ if and only if

$$||\nabla^2 f(x)|| \leq L, \quad \forall x \in \mathbf{R}^n.$$

Proof:

For one direction, if $f \in \mathcal{C}_L^{2,1}(\mathbf{R}^n)$, we have

- f is 2 times continuously differentiable in $\mathbf{R}^n$.
- its first derivative is Lipschitz continuous in $\mathbf{R}^n$ with the constant L:

$$||\nabla f(y) - \nabla f(x)|| \leq L||y - x||.$$

Thus, by the definition of derivative and Lipschitz condition, for any $s \in \mathbf{R}^n, \alpha \in \mathbf{R}^+$ we have

$$||(\int_0^\alpha \nabla^2 f(x + \tau s) d\tau) s|| = ||\nabla f(x + \alpha s) - \nabla f(x)|| \leq \alpha L||s||.$$

Taking the limit of α to 0, we obtain:

$$\lim_{\alpha \to 0} \frac{||\nabla f(x + \alpha s) - \nabla f(x)||}{\alpha} \leq L||s||,$$

$$\boxed{||\nabla^2 f(x)|| \leq L}.$$

For the other direction, by Mean value theorem, $\forall x, y \in \mathbf{R}^n$

$$\nabla f(y) = \nabla f(x) + \int_0^1 < \nabla^2 f(x + \tau(y - x)), y - x > d\tau$$

Thus, with the condition that the second derivative of f is bounded,

$$\begin{aligned} ||\nabla f(y) - \nabla f(x)|| &= ||\int_0^1 < \nabla^2 f(x+\tau(y-x)), y-x > d\tau|| \\ &\leq \int_0^1 || < \nabla^2 f(x+\tau(y-x)), y-x > ||d\tau \\ &\underset{\text{C.S.}}{\leq} \int_0^1 ||\nabla^2 f(x+\tau(y-x))||\ ||y-x||d\tau \\ &\leq \int_0^1 ||\nabla^2 f(x+\tau(y-x))||d\tau||y-x|| \leq L||y-x|| \end{aligned} \tag{1}$$

Corollary: Function f belongs to $\mathcal{C}_L^{p+1,p}(\mathbf{R}^n)$ if and only if

$$||\nabla^{p+1} f(x)|| \leq L, \quad \forall x \in \mathbf{R}^n.$$

Proof: For one direction, if $f \in \mathcal{C}_L^{p+1,p}(\mathbf{R}^n)$, for any $s \in \mathbf{R}^n, \alpha > 0$ we have

$$||(\int_0^\alpha \nabla^{p+1} f(x+\tau s)d\tau)s|| = ||\nabla^p f(x+\alpha s) - \nabla^p f(x)|| \leq \alpha L||s||.$$

Taking the limit of α to 0, we obtain:

$$\lim_{\alpha \to 0} \frac{||\nabla^p f(x+\alpha s) - \nabla^p f(x)||}{\alpha} \leq L||s||,$$

$$\boxed{||\nabla^{p+1} f(x)|| \leq L}.$$

For the other direction, $\forall x, y \in \mathbf{R}^n$

$$\nabla^p f(y) = \nabla^p f(x) + \int_0^1 < \nabla^{p+1} f(x+\tau(y-x)), y-x > d\tau$$

Thus,

$$\begin{aligned}
||\nabla^p f(y) - \nabla^p f(x)|| &= ||\int_0^1 < \nabla^{p+1} f(x + \tau(y-x)), y - x > d\tau|| \\
&\leq \int_0^1 || < \nabla^{p+1} f(x + \tau(y-x)), y - x > || d\tau \\
&\underset{\text{C.S.}}{\leq} \int_0^1 ||\nabla^{p+1} f(x + \tau(y-x))|| \; ||y - x|| d\tau \\
&\leq \int_0^1 ||\nabla^{p+1} f(x + \tau(y-x))|| d\tau ||y - x|| \leq L||y - x||
\end{aligned} \tag{2}$$

Remark:

- The Lipschitz condition actually comes many time from the Mean value theorem:

$$||f(y) - f(x)|| \leq ||\nabla f((1-\tau)x + \tau y)|| \; ||y - x||.$$

- ∇f is a vector, $\nabla^2 f$ is a Hessian matrix, $\nabla^p f$ is a tensor when $p \geq 3$.