Numerical Optimization with applications: Homework 07

104021601 林俊傑
104021602 吳彥儒
104021615 黃翊軒

December 14, 2016

Exercise 2 (Chapter 7)**.** *Show that the matrix* $\widehat{H}_{k+1} = (I - \frac{s_k y_k^T}{y_k^T s_k})$ *is singular.*

Proof. Consider $\widehat{H}_{k+1} s_k$, then we have

$$\begin{aligned} \widehat{H}_{k+1} s_k &= (I - \frac{s_k y_k^T}{y_k^T s_k}) s_k \\ &= s_k - \frac{s_k (y_k^T s_k)}{y_k^T s_k} \\ &= s_k - s_k \\ &= 0 \end{aligned}$$

Since $s_k = x_{k+1} - x_k \neq 0$, thus $\widehat{H}_{k+1}$ is singular. □

Exercise 5 (Chapter 10)**.** *Suppose that each residual function* r_j *and its gradient are Lipschitz continuous with Lipschitz constant L, that is ,*

$$\|r_j(x) - r_j(\widehat{x})\| \leq L\|x - \widehat{x}\|, \quad \| \bigtriangledown r_j(x) - \bigtriangledown r_j(\widehat{x})\| \leq L\|x - \widehat{x}\|$$

for all $j = 1, 2, ..., m$ *and all* $x, \widehat{x} \in \mathcal{D}$*, where* $\mathcal{D}$ *is a compact subset of* $\mathbb{R}^n$*. Assume also that the* r_j *are bounded on* $\mathcal{D}$*, that is there exist* $M > 0$ *such that* $|r_j(x)| \leq M$ *for all* $j = 1, 2, ..., m$ *and all* $x \in \mathcal{D}$*. Find Lipschitz constant for the Jacobian* J *and the gradient* $\bigtriangledown f$ *over* $\mathcal{D}$*.*

$$J(x) = \left[\frac{\partial r_j}{\partial x_i}\right]_{\substack{j=1,2,\ldots,m \\ i=1,2,\ldots,n}} = \begin{bmatrix} \bigtriangledown r_1(x)^T \\ \bigtriangledown r_2(x)^T \\ \vdots \\ \bigtriangledown r_m(x)^T \end{bmatrix}$$

$$\bigtriangledown f(x) = \Sigma_{j=1}^m r_j(x) \bigtriangledown r_j(x) = J(x)^T r(x)$$

Proof. Since all norms in $\mathbb{R}^n$ are equivalent.

$$\exists \alpha > 0 \quad \text{such that} \quad \|x\| \leq \alpha \|x\|_\infty \quad \forall x \in \mathbb{R}^n$$

We have,

$$\begin{aligned} \|J(x_1) - J(x_2)\| &= \max_{\|y\|=1} \|(J(x_1) - J(x_2))y\| \\ &= \max_{\|y\|=1} \left\| \begin{bmatrix} (\bigtriangledown r_1(x_1) - \bigtriangledown r_1(x_2))^T y \\ \vdots \\ (\bigtriangledown r_m(x_1) - \bigtriangledown r_m(x_2))^T y \end{bmatrix} \right\| \\ &\leq \max_{\|y\|=1} \alpha \left\| \begin{bmatrix} (\bigtriangledown r_1(x_1) - \bigtriangledown r_1(x_2))^T y \\ \vdots \\ (\bigtriangledown r_m(x_1) - \bigtriangledown r_m(x_2))^T y \end{bmatrix} \right\|_\infty \\ &= \alpha \max_{\|y\|=1} \max_{1 \leq j \leq m} |(\bigtriangledown r_j(x_1) - \bigtriangledown r_j(x_2))^T y| \\ &\leq \alpha \max_{\|y\|=1} \max_{1 \leq j \leq m} |(\bigtriangledown r_j(x_1) - \bigtriangledown r_j(x_2))| \, |y| \\ &\leq \alpha \max_{\|y\|=1} \max_{1 \leq j \leq m} L \|x_1 - x_2\| \; |y| \\ &= \alpha L \|x_1 - x_2\| \end{aligned}$$

We conclude that J is Lipschitz continuous with constant $\tilde{L} = \alpha L$.

On the other hand, Given $x, \tilde{x}$ in $\mathcal{D}$, we estimate

$$\begin{aligned}
\|\nabla f(x) - \nabla f(\tilde{x})\| &= \|J(x)^T r(x) - J(\tilde{x})^T r(\tilde{x})\| \\
&= \| \left[J(x)^T r(x) - J(\tilde{x})^T r(x)\right] + \left[J(\tilde{x})^T r(x) - J(\tilde{x})^T r(\tilde{x})\right] \| \\
&= \| \left(J(x)^T - J(\tilde{x})^T\right) r(x) + J(\tilde{x})^T \left(r(x) - r(\tilde{x})\right) \| \\
&\leq \|J(x)^T - J(\tilde{x})^T\| |r(x)| + \|J(\tilde{x})^T\| \|r(x) - r(\tilde{x})\| \\
&\leq M\alpha L \|x - \tilde{x}\| + M' L \|x - \tilde{x}\| \\
&= \mathcal{L} \|x - \tilde{x}\|
\end{aligned}$$

where $\mathcal{L} = M\alpha L + M'L$ and $\|J(\tilde{x})^T\|$ is bounded since it is Lipschitz continuous on a compact set $\mathcal{D}$. $\square$