Lecture 10

Thursday, 3 February 2022 1:54 PM

Convergence of Newton's Method

We need the following results:

① $A \in \mathbb{R}^{n\times n}$, $\lim_{k\to\infty} A^k = 0 \iff$ e.values satisfy $|\lambda_i(A)| < 1$, $i = 1, \dots, n$.

② The series of $n\times n$ matrices
$$I_n + A + A^2 + \dots + A^k + \dots$$
converges iff $\lim_{k\to\infty} A^k = 0$. Sum of the series is
$$(I_n - A)^{-1}.$$

③ Let $A : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be an $n\times n$ matrix valued function that is continuous at ξ_0 $\left(\lim_{\|\xi - \xi_0\|\to 0} \|A(\xi) - A(\xi_0)\| = 0\right)$

Proof later

If $A(\xi_0)^{-1}$ exists, then $A(\xi)^{-1}$ exists for ξ is sufficiently close to ξ_0 and $A(\cdot)^{-1}$ is continuous at ξ_0.

Convergence of Newton's Method

$\boxed{\text{Context}} \to \min_{x\in\mathbb{R}^n} f(x)$

Assumptions:
- (i) $f \in C^3$
- (ii) $x^* \in \mathbb{R}^n$ is such that $\nabla f(x^*) = 0$
- (iii) $F(x^*)$ is invertible.
- (iv) x^0 is sufficiently close to x^*.

$\left(x^{(k+1)} = x^{(k)} - F(x^{(k)})^{-1}\nabla f(x^{(k)})\right)$

Conclusions:
- Newton's method is well-defined for all k.
- Newton's method converges to x^* with an order of convergence at least 2.

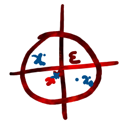
$\boxed{\text{PROOF}}$: Consider the Taylor expansion of ∇f.

$$\nabla f(x) = \nabla f(x^0) + F(x^0)(x - x^0) + \underbrace{O(\|x - x^0\|^2)}_{\text{remainder term.}}$$

$$\nabla f(\underline{x}) = \nabla f(\underline{x}^0) + F(\underline{x}^0)(\underline{x} - \underline{x}^0) + \underbrace{O(\|\underline{x} - \underline{x}^0\|)}_{\text{remainder term.}}$$

$f \in C^3$, $F(\underline{x}^*)$ is invertible $\Rightarrow \exists\ \varepsilon > 0,\ c_1 > 0,\ c_2 > 0$ such that

for all $\underline{x}, \underline{x}^{(0)} \in B_\varepsilon(\underline{x}^*)$, $= \{\underline{x} : \|\underline{x} - \underline{x}^*\| \le \varepsilon\}$



$$\|\nabla f(\underline{x}) - \nabla f(\underline{x}^0) - F(\underline{x}^0)(\underline{x} - \underline{x}^0)\| \le c_1 \|\underline{x} - \underline{x}^0\|^2 \quad \text{— ①}$$

for all $\underline{x} \in B_\varepsilon(\underline{x}^*)$

$$\|F(\underline{x})^{-1}\| \le c_2 \quad \text{— ③}$$

[$f \in C^{(3)}$, $\underline{x}, \underline{x}^0 \in B_\varepsilon(\underline{x}^*)$]
Cont. fn. in closed ball is bounded

Newton's method

$$\underline{x}^{(1)} = \underline{x}^{(0)} - F(\underline{x}^{(0)})^{-1} \nabla f(\underline{x}^{(0)})$$

$$\underline{x}^{(1)} - \underline{x}^* = \underline{x}^{(0)} - \underline{x}^* - F(\underline{x}^{(0)})^{-1} \nabla f(\underline{x}^{(0)})$$

$$\|\underline{x}^{(1)} - \underline{x}^*\| = \|\underline{x}^{(0)} - \underline{x}^* - F(\underline{x}^{(0)})^{-1} \nabla f(\underline{x}^{(0)})\|$$

$$= \|F(\underline{x}^0)^{-1} F(\underline{x}^0)(\underline{x}^{(0)} - \underline{x}^*) - F(\underline{x}^{(0)})^{-1} \nabla f(\underline{x}^{(0)})\|$$

$$\le \|F(\underline{x}^{(0)})\|^{-1} \|F(\underline{x}^{(0)})(\underline{x}^{(0)} - \underline{x}^*) - \nabla f(\underline{x}^{(0)})\|$$

$$\le c_2 c_1 \|\underline{x}^* - \underline{x}^{(0)}\|^2$$

In ①, choose $\underline{x} = \underline{x}^*$ (centre)

$$\|\underbrace{\nabla f(x^*)}_{=0} - \nabla f(\underline{x}^0) + F(\underline{x}^0)(\underline{x}^0 - \underline{x}^*)\| \le c_1 \|\underline{x}^* - \underline{x}^{(0)}\|^2$$

$$\|\underline{x}^{(1)} - \underline{x}^*\| \le c_1 c_2 \|\underline{x}^{(0)} - \underline{x}^*\|^2 \quad \text{— (*)}$$

Suppose $\underline{x}^{(0)}$ is such that $\|\underline{x}^{(0)} - \underline{x}^*\| \le \dfrac{\alpha}{c_1 c_2}$ $\quad 0 < \alpha < 1$

$$\|\underline{x}^{(1)} - \underline{x}^*\| \le \alpha \|\underline{x}^{(0)} - \underline{x}^*\| \quad \text{— (**)}$$

Induction : $\|\underline{x}^{(k+1)} - \underline{x}^*\| \le c_1 c_2 \|\underline{x}^{(k)} - \underline{x}^*\|^2$

$$\|\underline{x}^{(k+1)} - \underline{x}^*\| \le \alpha \|\underline{x}^{(k)} - \underline{x}^*\| \le \alpha^2 \|\underline{x}^{(k-1)} - \underline{x}^{(*)}\|$$

Induction : $\| \underline{x}^{(1)} - \underline{x}^{*} \| = c_1 c_2 \ldots$

$$\| \underline{x}^{(k+1)} - \underline{x}^{*} \| \le \alpha \| \underline{x}^{(k)} - \underline{x}^{*} \| \le \alpha^2 \| x^{(k-1)} - x^{(*)} \|$$

$$\vdots$$

$$\le \alpha^{k+1} \| \underline{x}^{(0)} - \underline{x}^{*} \|$$

Hence $\lim_{k \to \infty} \| \underline{x}^{k+1} - \underline{x}^{*} \| = 0$

$\{ \underline{x}^{(k)} \} \longrightarrow \underline{x}^{*}$.

Order of convergence is at least 2 since

$$\| \underline{x}^{k+1} - \underline{x}^{*} \| \le c_1 c_2 \| \underline{x}^{(k)} - \underline{x}^{*} \|^2$$

$$= O (\| x^{(k)} - \underline{x}^{*} \|^2).$$

May not be a descent algorithm! $\rightarrow \underbrace{f (\underline{x}^{(k+1)}) \ge f (\underline{x}^{(k)})}$.

To guarantee that Newton's method is indeed a descent,

$$d^{(k)} = x^{(k+1)} - \underline{x}^{(k)}$$

Exercise

$$\left[\begin{array}{l} \nabla f (x^{(k)}) \ne 0 \\ F (\underline{x}^{(k)}) > 0 \end{array} \right\}$$

$$x^{(k+1)} = x^{(k)} + \alpha d^{(k)}$$

Then we can show that $\exists \bar{\alpha} > 0$ such that

$$f (\underline{x}^{(k)} + \alpha d^{(k)}) < f (\underline{x}^{(k)}) \quad \forall \alpha \in (0, \bar{\alpha})$$

Modified Newton's method

$$\underline{x}^{(k+1)} = x^{(k)} - \alpha_k \underbrace{F (\underline{x}^k)^{-1}}_{(>0)} \underbrace{\nabla f (\underline{x}^k)}_{(\ne 0)}$$

$$\alpha_k = \arg \min_{\alpha \ge 0} \left[f (x^{(k)} - \alpha F (\underline{x}^k)^{-1} \nabla f (x^k)) \right]$$

Proof of ③ <u>Step 1</u> $A(\xi)^{-1}$ exists for ξ sufficiently close to $\underline{\xi}_0$.

<u>Step 2</u> A^{-1} is continuous at ξ_0.

Step 1 $A(\xi) = A(\xi_0) - A(\xi_0) + A(\xi)$.

Step 1.

$$A(\xi) = A(\xi_0) - A(\xi_0) + A(\xi).$$
$$= A(\xi_0)\left[I_n - I_n + A(\xi_0)^{-1} A(\xi)\right]$$
$$= A(\xi_0)\left[I_n - A(\xi_0)^{-1} A(\xi_0) + A(\xi_0)^{-1} A(\xi)\right]$$
$$= A(\xi_0)\left[I_n - \underbrace{A(\xi_0)^{-1}\left(A(\xi_0) - A(\xi)\right)}_{K(\xi)}\right]$$

$$A(\xi) = A(\xi_0)\left[I_n - K(\xi)\right]$$
$$\boxed{A(\xi)^{-1} = \left(I_n - K(\xi)\right)^{-1} A(\xi_0)^{-1}}$$

$$\|K(\xi)\| \leq \|A(\xi_0)^{-1}\| \, \underbrace{\|A(\xi_0) - A(\xi)\|}_{A \text{ is continuous at } \xi_0.}$$

$$\boxed{\lim_{\|\xi - \xi_0\| \to 0} \|K(\xi)\| = 0.}$$

$$\lim_{\|\xi - \xi_0\| \to 0} \|A(\xi_0) - A(\xi)\| = 0$$

$$\|A(\xi_0) - A(\xi)\| \leq \frac{\theta}{\|A(\xi_0)^{-1}\|}$$
$0 < \theta < 1$

for ξ_0 suff. close to ξ.

$$\|K(\xi)\| \leq \theta < 1$$
for ξ sufficiently close to ξ_0.

$\Rightarrow \left(I_n - K(\xi)\right)^{-1}$ exists. ✓ $\rightarrow$

Hence $A(\xi)^{-1}$ exists for ξ sufficiently close to ξ_0

Step 2 : Continuity of A^{-1}.

$$\|A(\xi)^{-1} - A(\xi_0)^{-1}\| = \|\left(I_n - K(\xi)\right)^{-1} \left(A(\xi_0)\right)^{-1} - A(\xi_0)^{-1}\|$$
$$= \|\left(\left(I_n - K(\xi)\right)^{-1} - I_n\right) A(\xi_0)^{-1}\|.$$
$$\leq \|\left(I_n - K(\xi)\right)^{-1} - I_n\| \, \|A(\xi_0)^{-1}\|$$

$^{-1}$ ... ^{2} ... $K^3(\xi) + \cdots$

$$\le \| (I_n - K(\xi))^{-1} - I_n \| \; \| A(\xi_0) \|$$

$$(I_n - K(\xi))^{-1} - I_n = K(\xi) + K^2(\xi) + K^3(\xi) + \cdots$$

$$= K(\xi) \left[I_n + K(\xi) + K^2(\xi) + \cdots \right].$$

$$\| (I_n - K(\xi))^{-1} - I_n \| \le \frac{\| K(\xi) \|}{1 - \| K(\xi) \|} \qquad \| K(\xi) \| < 1.$$

$$\| A(\xi)^{-1} - A(\xi_0)^{-1} \| \le \frac{\| K(\xi) \|}{1 - \| K(\xi) \|} \| A(\xi_0) \|^{-1}$$

$\because \lim_{\|\xi - \xi_0\| \to 0} \| K(\xi) \| = 0$, $\lim_{\|\xi - \xi_0\| \to 0} \| A(\xi)^{-1} - A(\xi_0)^{-1} \| = 0$

A^{-1} is continuous at ξ_0.