Background for the Convergence Analysis of Newton's Method for Systems

October 30, 2018

Given $f : \mathbb{R} \to \mathbb{R}$, $x, p \in \mathbb{R}$, define

$$\phi(\alpha) \ := \ f(x + \alpha p).$$

Fundamental theorem

$$\phi(1) = \phi(0) + \int_0^1 \phi'(t)\, dt$$

Given $f : \mathbb{R} \to \mathbb{R}$, $x, p \in \mathbb{R}$, define

$$\phi(\alpha) \ := \ f(x + \alpha p).$$

Fundamental theorem

$$\phi(1) = \phi(0) + \int_0^1 \phi'(t)\, dt$$

Taylor's theorem with integral remainder

$$\mathbf{f(x + p) = f(x) + \int_0^1 f'(x + tp) p \, dt}$$

Given $f : \mathbb{R}^n \to \mathbb{R}^n$, $x, p \in \mathbb{R}^n$, define

$$\phi_j(\alpha) \;:=\; f_j(x + \alpha p).$$

Fundamental theorem

$$\phi_j(1) = \phi_j(0) + \int_0^1 \phi_j'(t)\, dt$$

Equivalently,

$$f_j(x+p) = f_j(x) + \int_0^1 \nabla f_j(x+tp)^T p\, dt$$

Given $f : \mathbb{R}^n \to \mathbb{R}^n$, $x, p \in \mathbb{R}^n$, define

$$\phi_j(\alpha) \ := \ f_j(x + \alpha p).$$

Fundamental theorem

$$\phi_j(1) = \phi_j(0) + \int_0^1 \phi_j'(t)\, dt$$

Equivalently,

$$\begin{bmatrix} f_1(x+p) \\ f_2(x+p) \\ \vdots \\ f_n(x+p) \end{bmatrix} = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix} + \int_0^1 \begin{bmatrix} \nabla f_1(x+tp)^T p \\ \nabla f_2(x+tp)^T p \\ \vdots \\ \nabla f_n(x+tp)^T p \end{bmatrix} dt$$

Given $f : \mathbb{R}^n \to \mathbb{R}^n$, $x, p \in \mathbb{R}^n$, define

$$\phi_j(\alpha) \;:=\; f_j(x + \alpha p).$$

Fundamental theorem

$$\phi_j(1) = \phi_j(0) + \int_0^1 \phi_j'(t)\, dt$$

Equivalently,

$$\begin{bmatrix} f_1(x+p) \\ f_2(x+p) \\ \vdots \\ f_n(x+p) \end{bmatrix} = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix} + \int_0^1 \begin{bmatrix} \nabla f_1(x+tp)^T \\ \nabla f_2(x+tp)^T \\ \vdots \\ \nabla f_n(x+tp)^T \end{bmatrix} p\, dt$$

Given $f : \mathbb{R}^n \to \mathbb{R}^n$, $x, p \in \mathbb{R}^n$, define

$$\phi_j(\alpha) \; := \; f_j(x + \alpha p).$$

Fundamental theorem

$$\phi_j(1) = \phi_j(0) + \int_0^1 \phi_j'(t)\, dt$$

Taylor's theorem with integral remainder

$$\mathbf{f}(\mathbf{x} + \mathbf{p}) = \mathbf{f}(\mathbf{x}) + \int_{\mathbf{0}}^{\mathbf{1}} \mathbf{J_f}(\mathbf{x} + \mathbf{t}\mathbf{p})\mathbf{p}\, \mathbf{dt}$$

Definition 0.1 (Order of Convergence)**.** *Let* $\{x^{(k)}\}$ *be a sequence in* $\mathbb{R}^n$ *that converges to* x_**. Suppose also there exists a sequence* $\{\varepsilon_k\}$ *in* $\mathbb{R}$ *such that*

- $\|x^{(k)} - x_*\|_\infty \leq \varepsilon_k$*, and*
- $\lim_{k\to\infty} \varepsilon_{k+1}/\varepsilon_k^p = \mu$ *for some* $\mu, p \in \mathbb{R}$*,* $\mu > 0,\ p > 1.$

1. *We say* $\{x^{(k)}\}$ *converges to* x_* *with order at least* p*.*
2. *If* $\varepsilon_k = \|x^{(k)} - x_*\|_\infty$*,* $\{x^{(k)}\}$ *converges to* x_* *with order* p*.*
3. *If* $p = 2$*,* $\{x^{(k)}\}$ *converges to* x_* *at least quadratically.*
4. *If* $p = 2$ *and* $\varepsilon_k = \|x^{(k)} - x_*\|_\infty$*,* $\{x^{(k)}\}$ *converges to* x_* *quadratically.*