Regularity, estimates, and convergence of algorithms

Asen L. Dontchev

Mathematical Reviews and the University of Michigan

Supported by NSF Grant DMS-1008341

Classical Inverse Function Theorem

Inverse Function Theorem.

Let $f : X \to Y$ be C^1 near $\bar{x}$. Then FAE:

(a) The derivative mapping $Df(\bar{x})$ is invertible;

(b) f^{-1} has a C^1 single-valued localization around $f(\bar{x})$ for $\bar{x}$.

Lipschitz Continuous Functions

Definition

A function $f : X \to Y$ is **Lipschitz continuous** around $\bar{x}$ if $\exists \kappa > 0$ such that $\|f(x) - f(x')\| \leq \kappa \|x - x'\|$ for x, x' near $\bar{x}$.

Lipschitz modulus: $\operatorname{lip}(f; \bar{x})$ = limiting value of κ as $x, x'' \to \bar{x}$

Inverse Function Theorem for Lipschitz Functions.

Let $f, h : X \to Y$ and let

$$\operatorname{lip}(f - h; \bar{x}) = 0 \quad \text{and} \quad \bar{y} = f(\bar{x}) = h(\bar{x}).$$

Then FAE:

(a) f^{-1} has a Lipschitz continuous single-valued localization around $\bar{y}$ for $\bar{x}$;

(b) h^{-1} has a Lipschitz continuous single-valued localization around $\bar{y}$ for $\bar{x}$.

Strong Regularity (Robinson 1980)

Definition.

A mapping $F : X \rightrightarrows Y$ is said to be **strongly (metrically) regular at $\bar{x}$ for $\bar{y}$** when $\bar{y} \in F(\bar{x})$ and F^{-1} has a Lipschitz continuous single-valued localization around $\bar{y}$ for $\bar{x}$;

Inverse Function Theorem for Lipschitz Functions.

Let $f, h : X \to Y$ and let

$$\mathrm{lip}(f - h; \bar{x}) = 0 \quad \text{and} \quad \bar{y} = f(\bar{x}) = h(\bar{x}).$$

Then FAE:

(a) f is strongly regular at $\bar{x}$ for $\bar{y}$;

(b) h is strongly regular at $\bar{x}$ for $\bar{y}$.

Extension to Set-Valued Mappings

Theorem.

Let $f, h : X \to Y$, $F : X \rightrightarrows Y$ and let

$$\operatorname{lip}(f - h; \bar{x}) = 0, \ \ f(\bar{x}) = h(\bar{x}) = \bar{y} \ \text{ and } \ \bar{z} \in F(\bar{x}).$$

Then FAE:

(a) $f + F$ is strongly regular at $\bar{x}$ for $\bar{y} + \bar{z}$;

(b) $h + F$ is strongly regular at $\bar{x}$ for $\bar{y} + \bar{z}$.

Metric Regularity

Definition.

A mapping $F : X \rightrightarrows Y$ is said to be **metrically regular at $\bar{x}$ for $\bar{y}$** when $\bar{y} \in F(\bar{x})$ and there is a constant $\kappa \geq 0$ together with neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$d(x, F^{-1}(y)) \leq \kappa d(y, F(x)) \quad \text{for all } (x, y) \in U \times V.$$

The infimum of κ is the regularity modulus $\mathrm{reg}(F; \bar{x} \,|\, \bar{y})$

Theorem (Banach open mapping theorem).

A mapping $A \in \mathcal{L}(X, Y)$ is metrically regular (anywhere) if and only of it is surjective.

Lyusternik-Graves theorem

Theorem.

Let X and Y be Banach spaces and let $f : X \to Y$ be continuously Fréchet differentiable around $\bar{x}$. Then FAE:

(i) the derivative mapping $Df(\bar{x})$ is surjective (or equivalently, metrically regular);

(ii) f is metrically regular at $\bar{x}$ for $f(\bar{x})$.

Theorem.

Let $f, h : X \to Y$, let $F : X \rightrightarrows Y$, let

$$\mathrm{lip}(f - h; \bar{x}) = 0, f(\bar{x}) = h(\bar{x}) = \bar{y} \text{ and } \bar{z} \in F(\bar{x}),$$

and let F have locally closed graph at $(\bar{x}, \bar{z})$. Then FAE:

(a) $f + F$ is metrically regular at $\bar{x}$ for $\bar{y} + \bar{z}$;

(b) $h + F$ is metrically regular at $\bar{x}$ for $\bar{y} + \bar{z}$.

Radius Theorem

Radius Theorem [A.D., A. Lewis, R. T. Rockafellar].

Consider a mapping $F : X \rightarrow Y$ with closed graph and let F be metrically regular at $\bar{x}$ for $\bar{y}$. Then

$$\inf_{f:X\to Y} \left\{ \operatorname{lip}(f;\bar{x}) \,\middle|\, F+f \text{ not met. reg. at } \bar{x} \text{ for } \bar{y}+f(\bar{x}) \right\} \geq \frac{1}{\operatorname{reg}(F;\bar{x}\,|\,\bar{y})}.$$

If X and Y are finite-dimensional then $\leq$ becomes $=$ and then the infimum is the same when restricted to linear mappings of rank one.

$\operatorname{reg}(F;\bar{x}\,|\,\bar{y})$ absolute condition number

Radius Theorem for strong metric regularity

Metric subregularity

Definition.

A mapping $F : X \rightrightarrows Y$ is said to be **metrically subregular at $\bar{x}$ for $\bar{y}$** when $\bar{y} \in F(\bar{x})$ and there is a constant $\kappa \geq 0$ together with a neighborhood U of $\bar{x}$ such that

$$d(x, F^{-1}(\bar{y})) \leq \kappa d(\bar{y}, F(x)) \quad \text{for all } x \in U.$$

Error bounds, deriving the Lagrange multiplier rule

Does not obey the paradigm of the implicit function theorem: the radius of metric subregularity is zero.

Strong metric subregularity

Definition.

A mapping $F : X \rightrightarrows Y$ is said to be **strongly metrically subregular at $\bar{x}$ for $\bar{y}$** when $\bar{y} \in F(\bar{x})$ and there is a constant $\kappa \geq 0$ together with a neighborhood U of $\bar{x}$ such that

$$\|x - \bar{x}\| \leq \kappa d(\bar{y}, F(x)) \quad \text{for all } x \in U.$$

Theorem.

Let $f, h : X \to Y$, let $F : X \rightrightarrows Y$, let

$$\operatorname{lip}(f - h; \bar{x}) = 0, \bar{y} = f(\bar{x}) = h(\bar{x}) \text{ and } \bar{z} \in F(\bar{x})$$

and let F have locally closed graph at $(\bar{x}, \bar{z})$. Then FAE:

(a) $f + F$ is strongly metrically subregular at $\bar{x}$ for $\bar{y} + \bar{z}$;

(b) $h + F$ is strongly metrically subregular at $\bar{x}$ for $\bar{y} + \bar{z}$.

Regularity in convex optimization

$$\text{minimize} \quad g(x) - \langle p, x\rangle \qquad \text{over} \quad x \in C,$$

$g : \mathbf{R}^n \to \mathbf{R}$ convex and C^2, $p \in \mathbf{R}^n$ parameter, and C convex polyhedral.

First-order optimality condition $\nabla g(x) + N_C(x) \ni p$

The mapping $\nabla g + N_C$ is strongly metrically subregular at $\bar{x}$ for $\bar{p}$ if and only if the standard second-order sufficient condition holds at $\bar{x}$ for $\bar{p}$: $\langle \nabla^2 g(\bar{x})u, u\rangle > 0$ for all nonzero u in the critical cone $K_C(\bar{x}, \bar{p} - \nabla g(\bar{x}))$.

The mapping $\nabla g + N_C$ is metrically subregular at $\bar{x}$ for $\bar{p}$ if and only if it is strongly metrically regular, which is equivalent to the strong second-order sufficient condition at $\bar{x}$ for $\bar{p}$:
$\langle \nabla^2 g(\bar{x})u, u\rangle > 0$ for all nonzero
$u \in K_C(\bar{x}, \bar{p} - \nabla g(\bar{x})) - K_C(\bar{x}, \bar{p} - \nabla g(\bar{x}))$.

Newton's method for variational inequalities

Variational inequality: $f : \boldsymbol{R}^n \to \boldsymbol{R}^n$, $f \in C^1$, $C \subset \boldsymbol{R}^n$ closed and convex

$$\langle f(x), y - x \rangle \leq 0 \text{ for all } y \in C$$

Equivalently,

$$f(x) + N_C(x) \ni 0$$

Newton's method

$$f(x_k) + Df(x_k)(x_{k+1} - x_k) + N_C(x_{k+1}) \ni 0$$

Convergence of Newton's method

Theorem.

Consider Newton's method for a function f which is continuously differentiable near $\bar{x}$ and such that $\mathrm{lip}(Df;\bar{x}) < \infty$. Assume that the mapping $f + N_C$ is strongly metrically regular at $\bar{x}$ for 0. Then there exists a neighborhood O of $\bar{x}$ such that, for any $x_0 \in O$, there exists a unique sequence $\{x_k\}$ generated by the method which is quadratically convergent to $\bar{x}$.

metric regularity: ... there exists a sequence generated by the method which converges quadratically.

strong metric subregularity: ... any sequence generated by the method which stays in O converges quadratically.

Perturbed Newton's method

$$X \times P \ni (u,p) \mapsto \Xi(u,p) = \Big\{ \{x_k\} \in l_\infty(X) \,\Big|\, x_0 = u,$$

$$p \in f(x_k) + Df(x_k)(x_{k+1} - x_k) + N_C(x_{k+1})), \; \forall k = 1, 2, \dots \Big\},$$

Theorem.

Let $\operatorname{lip}(Df; \bar{x}) < \infty$. Then FAE:

(i) the mapping $f + N_C$ is strongly metrically regular at 0 for $\bar{x}$;

(ii) the mapping Ξ has a Lipschitz continuous single-valued localization around $(\bar{x}, 0)$ for $\{\bar{x}\}$ (i.e., Ξ^{-1} is strongly metrically regular) each value of which is a sequence which is quadratically convergent with the same constant.

Inexact Newton's method

R. Dembo, S. C. Eisenstat, T. Steihaug, Inexact Newton methods, SIAM J. Numer. Anal. 19 (1982), no. 2, 400–408.

$f : \boldsymbol{R}^n \to \boldsymbol{R}$, $f(\bar{x}) = 0$, $Df(\bar{x})$ nonsingular

$$\|f(x_k) + Df(x_k)(x_{k+1} - x_k)\| \leq \eta \|f(x_k)\|^2$$

If x_0 is close to $\bar{x}$, then every sequence $\{x_k\}$ generated by the method converges to $\bar{x}$ quadratically;

Inexact Newton's method for VI

Theorem (convergence under strong metric regularity).

Let the mapping $f + N_C$ be strongly metrically regular at $\bar{x}$ for 0. Let $\varphi(x) = P_C(f(x) - x) + x$ and consider the inexact Newton iteration

$$d(0, f(x_k) + Df(x_k)(x_{k+1} - x_k) + N_C(x_{k+1})) \leq \eta \|\varphi(x_k)\|^2$$

Then there exists a neighborhood O of $\bar{x}$ and $C > 0$ such that for any $x_0 \in O$ there exists a sequence $\{x_k\}$ generated by the method and every such sequence converges to $\bar{x}$ quadratically.

metric regularity: ... there exists a sequence generated by the method which converges quadratically.

strong metric subregularity: ... any sequence generated by the method which stays in O converges quadratically.

Dennis-Moré Theorem

Quasi-Newton method for solving $f(x) = 0$:

$$f(x_k) + B_k(x_{k+1} - x_k) = 0,$$

where $f : \boldsymbol{R}^n \to \boldsymbol{R}^n$ and B_k is a sequence of matrices.
Let $s_k = x_{k+1} - x_k$, $e_k = x_k - \bar{x}$, $E_k = B_k - Df(\bar{x})$. Recall that $\{x_k\}$ converges superlinearly when $\|e_{k+1}\|/\|e_k\| \to 0$.

Theorem [Dennis-Moré, 1974].

Suppose that f is differentiable near a zero $\bar{x}$, the derivative Df is continuous at $\bar{x}$ and $Df(\bar{x})$ is nonsingular. Let $\{B_k\}$ be a sequence of nonsingular matrices. Consider a sequence $\{x_k\}$ generated by the method for some starting point x_0 near $\bar{x}$. Then $x_k \to \bar{x}$ superlinearly if and only if

$$x_k \to \bar{x} \quad \text{and} \quad \lim_{k\to\infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0.$$

Dennis-Moré Theorem for VI

Theorem.

Let C be convex polyhedral and suppose that $f + N_C$ is strongly metrically subregular at $\bar{x}$ for 0. Consider a sequence $\{x_k\}$ generated by

$$f(x_k) + B_k(x_{k+1} - x_k) + N_C(x_{k+1}) \ni 0,$$

for some starting point x_0 near $\bar{x}$. If $x_k \to \bar{x}$ superlinearly, then

$$\lim_{k\to\infty} \frac{d(0, E_k s_k + N_K(e_{k+1}))}{\|s_k\|} = 0 \quad (K \text{ is the critical cone at } \bar{x}).$$

Conversely, is $\{x_k\}$ is such that

$$x_k \to \bar{x} \quad \text{and} \quad \lim_{k\to\infty} \frac{\|E_k s_k\|}{\|s_k\|} = 0,$$

then $x_k \to \bar{x}$ superlinearly.

Conjecture

The condition

$$x_k \to \bar{x} \text{ and } \lim_{k\to\infty} \frac{d(0, E_k s_k + N_K(e_{k+1}))}{\|s_k\|} = 0.$$

is a necessary and sufficient condition for superlinear convergence.