Inverse Function Theorem

Modifications of the standard case

Bernadett Stadler

November 17^{th}, 2015

Introduction

Theorem (The classical Inverse Function Theorem)

Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in a neighborhood of a point $\overline{x}$ and let $\overline{y} := f(\overline{x})$. If $\nabla f(\overline{x})$ is nonsingular, then f^{-1} has a single-valued localization s around $\overline{y}$ for $\overline{x}$. Moreover, the function s is continuously differentiable in a neighborhood V of $\overline{y}$, and its Jacobian satisfies $\nabla s(y) = \nabla f(s(y))^{-1}$ for every $y \in V$.

Modifications of the classical case:

1. depart from differentiability
2. $f : \mathbb{R}^n \to \mathbb{R}^m$ with $m \leq n$

Outline

1. Introduction
2. Modification 1: Non-differentiable functions
   - Inverse Function Theorem Beyond Differentiation
3. Modification 2: Selections of Multi-Valued Inverses
   - Inverse Selections for $m \leq n$
4. Modification 3: Selections from non-differentiable functions
   - Inverse Selections from First-Order Approximation

Modification 1: Depart from Differentiation

Approach:

- approximation of non-differentiable f by another function h around $\overline{x}$
- f differentiable $\Rightarrow$ can take linearization $h(x) = f(\overline{x}) + \nabla f(\overline{x})(x - \overline{x})$
- differentiability corresponds to $clm(e; \overline{x}) = 0$
  strict differentiability to $lip(e; \overline{x}) = 0$
  with $e(x) = f(x) - [f(\overline{x}) - A(x - \overline{x})]$
- key idea: conditions can be applied even if h is not a linearization

Outline

Preliminary Definitions

Definition (First-Order Approximations and Estimators of Functions)

Assumptions:

- $f : \mathbb{R}^n \to \mathbb{R}^m$ and $h : \mathbb{R}^n \to \mathbb{R}^m$ with $h(\overline{x}) = f(\overline{x})$
- $\overline{x} \in \text{int dom } f$ and $\overline{x} \in \text{int dom } h$
- $e(x) = f(x) - h(x)$

Then h is called

1. a **first-oder approximation** to f at $\overline{x}$: $clm(e; \overline{x}) = 0$
2. a **strict first-order approximation**: $lip(e; \overline{x}) = 0$
3. an **estimator** of f at $\overline{x}$: $clm(e; \overline{x}) \leq \mu < \infty$
4. a **strict estimator**: $lip(e; \overline{x}) \leq \mu < \infty$

Properties

- If q is a (strict) first-order approximation to h at $\overline{x}$ and h is a (strict) first-order approximation to f at $\overline{x}$, then q is a (strict) first-order approximation to f at $\overline{x}$.
- If f_1 and f_2 have (strict) first-order approximation h_1 and h_2 at $\overline{x}$, then $h_1 + h_2$ is a (strict) first-oder approximation of $f_1 + f_2$ at $\overline{x}$.
- If f has a (strict) first-oder approximation h at $\overline{x}$, then for any $\lambda \in \mathbb{R}$, λh is a (strict) first-order approximation of λf at $\overline{x}$.
- If h is a first-order approximation of f at $\overline{x}$, then $clm(f;\overline{x}) = clm(h;\overline{x})$.
- If h is a strict first-order approximation of f at $\overline{x}$, then $lip(f;\overline{x}) = lip(h;\overline{x})$.

Corollary (Composition of First-Order Approximations)

Assumptions:

- *h is a first-order approximaton of f at $\overline{x}$*
- $clm(h;\overline{x}) < \infty$
- *v is a first-order approximation of u which is Lipschitz continuous around $\overline{y}$*

Then $v \circ h$ is a first-order approximation of $u \circ f$ at $\overline{x}$.

Theorem (Inverse Function Theorem Beyond Differentiation)

Assumptions:

- $f : \mathbb{R}^n \to \mathbb{R}^n$ *with* $\overline{x} \in$ *int dom f*
- $h : \mathbb{R}^n \to \mathbb{R}^n$ *be a strict estimator of* f *at* $\overline{x}$ *with constant* μ
- h^{-1} *has a Lipschitz continuous single-valued localization* σ *around* $\overline{y}$
- $lip(\sigma, \overline{y}) \leq \kappa$ *for a* κ *such that* $\kappa\mu < 1$

Then f^{-1} *has a Lipschitz continuous single-valued localization s around* $\overline{y}$ *for* $\overline{x}$ *with* $lip(s; \overline{y}) \leq \frac{\kappa}{1-\kappa\mu}$.

Proof Part 1:

w.l.o.g. assume $\bar{x} = \bar{y} = 0$

Let $x^{k+1} := \sigma(y - e(x^k))$.
by induction: $x^k \in a\mathbb{B}$, $y - e(x^k) \in a\mathbb{B}$, $|x^k - x^{k-1}| \leq (\lambda \nu)^{k-1}|x^1 - x^0|$

Fix: $\lambda \in (\kappa, \infty)$ and $\nu \in (\mu, \kappa^{-1})$ such that $\lambda \nu < 1$
take a small enough such that:

- $y \mapsto h^{-1}(y) \cap a\mathbb{B}$ $y \in a\mathbb{B}$ is a localization of σ which is Lipschitz continuous with constant λ
- and $lip(e; \bar{x}) \leq \nu$ $\bar{x} \in a\mathbb{B}$

take α such that: $0 < \alpha < \frac{1}{4}a(1 - \lambda \nu)$
and define $b := \frac{\alpha}{4\lambda}$

Proof Part 2: Show the nonempty-valuedness of the localization of f^{-1}.

x^k is a Cauchy Sequence $\Rightarrow x^k \to x$ for $k \to \infty$
$x = \lim x^{k+1} = \lim \sigma(y - e(x^k)) = \sigma(y - e(x))$
$\Rightarrow x \in f^{-1}(y)$

from the Lipschitz continuity of σ and e we obtain: $|x| \leq \frac{\lambda b}{1-\lambda v}$

Thus $\forall y \in b\mathbb{B} \exists x \in f^{-1}(y)$ with $|x|$ satisfying the estimation above.
$\Rightarrow$ have shown the nonempty-valuedness of the localization s of f^{-1}:

$$s : y \mapsto f^{-1}(y) \cap \frac{\lambda b}{1-\lambda v}\mathbb{B}$$

Proof Part 3: s is single-valued and Lipschitz continuous.

take x, x' with $x \neq x'$ and a fix $y \in b\mathbb{B}$:
$0 < |x - x'| = |\sigma(y - e(x')) - \sigma(y - e(x))| \leq \lambda \nu |x - x'| < |x - x'|$
which is a contradiction $\Rightarrow$ s is single valued

take $y, y' \in b\mathbb{B}$:
$|s(y') - s(y)| = |\sigma(y' - e(x')) - \sigma(y - e(x))| \leq \lambda(\nu |s(y') - s(y)| + |y' - y|)$
$\Leftrightarrow |s(y') - s(y)| \leq \frac{\lambda}{1 - \lambda\nu} |y' - y|$

$$|s(y') - s(y)| \leq \frac{\kappa}{1 - \kappa\mu} |y' - y|$$

□

Example: Loss of Single-Valued Localization without strict differentiability

Let $\alpha \in (0, \infty)$ and $f(x) := \alpha x + g(x)$ with

$$g(x) = \begin{cases} x^2 sin(1/x) & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$$

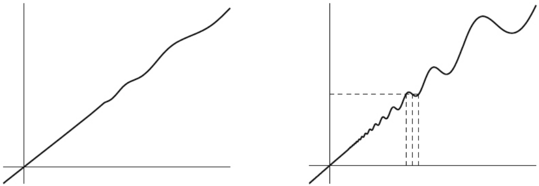

Figure: Graphs with $\alpha = 2$ and $\alpha = 0.5$

Special Case

If $\mu = 0$ (h is a strict first-order approximation) one can proof an equivalence emerges between f^{-1} and h^{-1}:

Theorem (Lipschitz Invertibility with First-Order Approximation)

Assumptions:

- $f : \mathbb{R}^n \to \mathbb{R}^n$
- $h : \mathbb{R}^n \to \mathbb{R}^n$ *is a strict first-order approximation of* f *at* $\overline{x}$
- *let* $\overline{y} = f(\overline{x}) = h(\overline{x})$

Then f^{-1} *has a Lipschitz continuous single-valued localization* s *around* $\overline{y}$ *for* $\overline{x}$ **if and only if** h^{-1} *has such a localization* σ *around* $\overline{y}$ *for* $\overline{x}$. *Furthermore* $lip(s;\overline{y}) = lip(\sigma;\overline{y})$ *and* σ *is a first-order approximation to* s *at* $\overline{y}$.

Modification 2

- $f : \mathbb{R}^n \to \mathbb{R}^m$ with $m \leq n$
- What can be said about f^{-1}?

Example: $f(x) = Ax + b$ with $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$

- $m < n$: $Ax + b = y$ has no solution or a continuum of solutions
  $\Rightarrow$ single-valued localizations does not exist
- A has full rank: $f^{-1}(y)$ is nonempty $\forall y$

Question: Do we really always need to assume that $m = n$ to get a single valued localization?

Outline

Preliminary Theorem

Theorem (Brouwer Invariance of Domain Theorem)

Assumptions:

- $O \subset \mathbb{R}^n$ *open*
- $f : O \to \mathbb{R}^m$ *continuous, for* $m \leq n$
- f^{-1} *single-valued on* $f(0)$

Then $f(O)$ *is open,* f^{-1} *continuous on* $f(O)$ *and* $m = n$.

- can conlude for $f : \mathbb{R}^n \to \mathbb{R}^m$:
  If $m < n \Rightarrow f^{-1}$ fails to have a single-valued localization
- however one can proof a different statement with the concept of selections

Selections

Definition (Selections)

Assumptions:

- $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a set $D \subset \operatorname{dom} F$
- $w : \mathbb{R}^n \to \mathbb{R}^m$ with $D \subset \operatorname{dom} w$ and $w(x) \in F(x)$

Then w is a **selection** of F on D.
If in addition $(\overline{x}, \overline{y}) \in \operatorname{gph} F$, D is a neigborhood of $\overline{x}$ and $w(\overline{x}) = \overline{y}$. Then w is a **local selection** of F around $\overline{x}$ for $\overline{y}$.

Selections

A selection of f^{-1} might provide a left or a right inverse of f.

- **left inverse** to f on D: a selection $l : \mathbb{R}^m \to \mathbb{R}^n$ of f^{-1} on $f(D)$ such that $l(f(x)) = x$
- **right inverse** to f on D: a selection $r : \mathbb{R}^m \to \mathbb{R}^n$ of f^{-1} on $f(D)$ such that $f(r(y)) = y$

Example: Linear mapping from $\mathbb{R}^m$ to $\mathbb{R}^n$ represented by a matrix $A \in \mathbb{R}^{m \times n}$ of full rank. If

- $m \leq n$: right inverse of A corresponds to $A^T(AA^T)^{-1}$
- $m \geq n$: left inverse of A corresponds to $(A^TA)^{-1}A^T$
- $m = n$: right and left inverse are equal to A^{-1}

Theorem (Inverse Selections for $m \leq n$)

Assumptions:

- $f : \mathbb{R}^n \to \mathbb{R}^m$ *with* $m \leq n$
- f *is* k *times continuously differentiable in a neighborhood of* $\overline{x}$
- $\nabla f(\overline{x})$ *has full rank* m

Then there exists a local selection s *of* f^{-1} *around* $\overline{y}$ *for* $\overline{x}$ *which is* k *times continuously differentiable in a neighborhood* V *of* $\overline{y}$ *and the Jacobian satisfies:*

$$\nabla s(\overline{y}) = A^T(AA^T)^{-1} \text{ with } A := \nabla f(\overline{x})$$

Special Case

Theorem (Strictly Differentiable Selections)

- $f : \mathbb{R}^n \to \mathbb{R}^m$ *with* $m \le n$
- f *is strictly differentiable with* $\nabla f(\overline{x})$ *is of full rank*

Then there exists a local selection s *of the inverse* f^{-1} *around* $\overline{y}$ *for* $\overline{x}$ *which is strictly differentiable at* $\overline{y}$ *and* $\nabla s(\overline{y}) = A^T(AA^T)^{-1}$.

Outline

Theorem (Inverse Selections from First-Order Approximation)

Assumptions:

- $f : \mathbb{R}^n \to \mathbb{R}^m$ *continuous around* $\overline{x}$ *with* $f(\overline{x}) = \overline{y}$
- $h : \mathbb{R}^n \to \mathbb{R}^m$ *first-order approximation of* f *at* $\overline{x}$ *which is continuous around* $\overline{x}$
- h^{-1} *has a Lipschitz continuous local selection* σ

Then f^{-1} *has a local selection* s *around* $\overline{y}$ *for* $\overline{x}$.
Furthermore σ *is a first-order approximation of* s *at* $\overline{y}$.

Preliminary Theorem

Theorem (Brouwer Fixed Point Theorem)

Assumptions:

- Q *compact and convex set in* $\mathbb{R}^n$
- $\phi : \mathbb{R}^n \to \mathbb{R}^n$ *continuous on* Q
- ϕ *maps* Q *into itself*

Then $\exists x \in Q : \phi(x) = x$.

Sketch of the Proof Part 1:

Show that f^{-1} has a local selections:

w.l.o.g $\overline{x} = \overline{y} = 0$

1. choose an a such that f and h are continuous on $a\mathbb{B}$
2. define: $\phi_y : x \mapsto \sigma(y - e(x)) \; \forall x \in a\mathbb{B}$

since e is continuous, and σ is Lipschitz continuous on $a\mathbb{B}$
$\Rightarrow \phi$ is continuous on $a\mathbb{B}$

with Brouwer's theorem one can conclude that:
$\exists x \in a\mathbb{B} : x = \sigma(y - e(x))$

$$h(x) = h(\sigma(y - e(x))) = y - e(x) \Leftrightarrow f(x) = y$$

Sketch of the Proof Part 2:

For each $y \in b\mathbb{B}$ pick one such fixed point x of the function ϕ_y and define $s(y) := x$.
$\Rightarrow s(y) = \sigma(y - e(s(y)))\ \forall y \in b\mathbb{B}$

$y = 0$ we set $s(0) = 0 \in f^{-1}(0)$

$\Rightarrow$ **s is a local selection of f^{-1} around 0 for 0**

For an arbitrary neighborhood U of 0 we found $b > 0$ such that $s(y) \in U\ \forall y \in b\mathbb{B}$.
$\Rightarrow$ **s is continuous in 0** □

Special Case

Theorem (Inverse Selections from Nonstrict Differentiability)

Assumptions:

- $f : \mathbb{R}^n \to \mathbb{R}^n$ *continuous in a neighborhood of* $\overline{x} \in \operatorname{int} \operatorname{dom} f$
- f *is differentiable at* $\overline{x}$ *with* $\nabla f(\overline{x})$ *nonsingular*

Then there exists a local selection of f^{-1} *around* $\overline{x}$ *which is continuous at* $\overline{y}$*. Moreover, every local selection* s *of* f^{-1} *around* $\overline{y}$ *which is continuous at* $\overline{y}$ *has the property that*

$$s \text{ is differentiable at } \overline{y} \text{ with } \nabla s(\overline{y}) = \nabla f(\overline{x})^{-1}.$$

Theorem (Inverse Function Theorem Beyond Differentiation)

Assumptions:

- $f : \mathbb{R}^n \to \mathbb{R}^n$ *with* $\overline{x} \in$ *int dom* f
- $h : \mathbb{R}^n \to \mathbb{R}^n$ *be a strict estimator of* f *at* $\overline{x}$ *with constant* μ
- h^{-1} *has a Lipschitz continuous single-valued localization* σ *around* $\overline{y}$
- $lip(\sigma, \overline{y}) \leq \kappa$ *for a* κ *such that* $\kappa\mu < 1$

Then f^{-1} *has a Lipschitz continuous single-valued localization* s *around* $\overline{y}$ *for* $\overline{x}$ *with* $lip(s; \overline{y}) \leq \frac{\kappa}{1-\kappa\mu}$.

Example: Inverse Multi-Valuedness if Differentiability is not Strict

Let $\alpha \in (0,\infty)$ and $f(x) := \alpha x + g(x)$ with

$$g(x) = \begin{cases} x^2 sin(1/x) & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$$

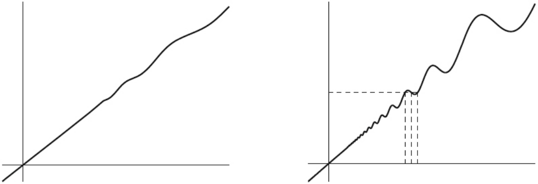

Figure: Graphs with $\alpha = 2$ and $\alpha = 0.5$

Summary

1. Depart from differentiation, approximate f
   Inverse function theorem beyond differentiation
   Special Case: Equivalence between f^{-1} and h^{-1}
2. $f : \mathbb{R}^n \to \mathbb{R}^m$ with $m \leq n$
   Brouwer's invariance of domain theorem
   Selections
   Inverse Selections
3. Combine 1 and 2
   Inverse Selections from first-order approximations
   Special Case: Inverse Selections from nonstrict differentiability