A Hierarchy of Existence Theorems for Nonlinear Equations

Andreas Frommer
Bergische Universität Wuppertal
Fachbereich Mathematik
frommer@math.uni-wuppertal.de
http://www.math.uni-wuppertal.de/SciComp

joint work with

Götz Alefeld
Gerhard Heindl
Jan Mayer

Contents

Contents

Contents

Contents

Back
Close

Contents

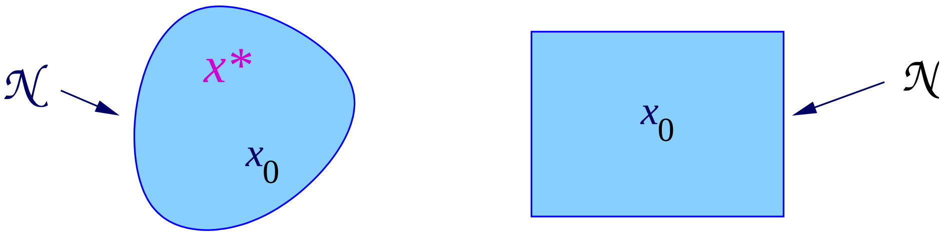
Contents

Contents

General setting

$f:\mathbb{R}^n \to \mathbb{R}^n$	Fréchet-differentiable,
$f'(x) \in \mathbb{R}^{n\times n}$	Jacobian
$x_0 \in \mathbb{R}^n$	approximate zero

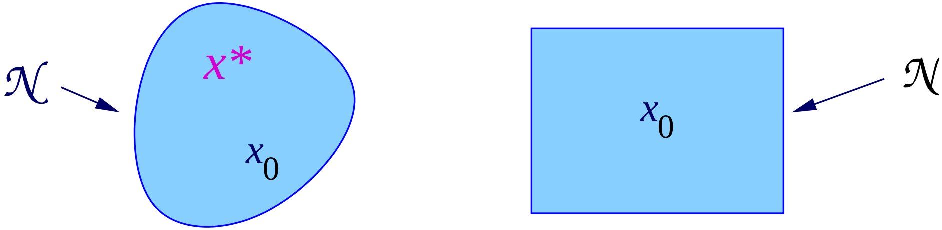
$\mathcal{N} \subseteq \mathbb{R}^n$, $x_0 \in \mathcal{N}$ „neighbourhood“



General setting

$f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$	Fréchet-differentiable,
$f'(x) \in \mathbb{R}^{n\times n}$	Jacobian
$x_0 \in \mathbb{R}^n$	approximate zero

$\mathcal{N} \subseteq \mathbb{R}^n$, $x_0 \in \mathcal{N}$ „neighbourhood“



Question: Is there a zero x^* of f in $\mathcal{N}$?
(conditions on $x_0, f, f', \mathcal{N}$ etc.)

Moore's existence test

$\mathcal{N} = [x] \subseteq \mathbb{R}^n$ interval vector

Theorem (Moore, 1977)
Let $x^0 \in [x]$. Assume

a) $\forall x \in [x] : f(x) - f(x^0) \in [Y](x - x^0)$
for some interval matrix $[Y]$

b) $x^0 - A \cdot f(x^0) + (I - A[Y])([x] - x^0) \subseteq [x]$
for some nonsingular matrix $A \in \mathbb{R}^{n \times n}$

Then there exists $x^* \in [x]$ s.t. $f(x^*) = 0$.

Moore's existence test

$\mathcal{N} = [x] \subseteq \mathbb{R}^n$ interval vector

Theorem (Moore, 1977)
Let $x^0 \in [x]$. Assume

a) $\forall x \in [x] : f(x) - f(x^0) \in [Y](x - x^0)$
for some interval matrix $[Y]$

b) $x^0 - A \cdot f(x^0) + (I - A[Y])([x] - x^0) \subseteq [x]$
for some nonsingular matrix $A \in \mathbb{R}^{n \times n}$

Then there exists $x^* \in [x]$ s.t. $f(x^*) = 0$.

Choices for $[Y]$:

a) $[Y] = \square\{f'(x), x \in [x]\}$ interval hull of derivative

b) $[Y] = \square\{\delta f(x, x_0), x \in [x]\}$ interval hull of slopes
$f(x) - f(x_0) = \delta f(x, x_0)(x - x_0)$

Newton-Kantorovich

$\mathcal{N}$ = norm ball, $\mathcal{N} = B(x_0, \varrho) = \{x : \|x - x_0\| \leq \varrho\}$

Theorem (Kantorovich, 1948, Heindl, Deuflhard, 1980)
Assume that in some open convex D

Newton-Kantorovich

$\mathcal{N}$ = norm ball, $\mathcal{N} = B(x_0, \varrho) = \{x : \|x - x_0\| \leq \varrho\}$

Theorem (Kantorovich, 1948, Heindl, Deuflhard, 1980)
Assume that in some open convex D

a) $\|f'(x_0)^{-1}\| \leq \beta$,
$\|f'(x_0)^{-1} f(x_0)\| \leq \eta$

b) $\|f'(u) - f'(v)\| \leq \kappa \cdot \|u - v\|$

c) $\omega = \beta\kappa,\ h = \eta\omega \leq \frac{1}{2}$

(standard form)

Newton-Kantorovich

$\mathcal{N}$ = norm ball, $\mathcal{N} = B(x_0, \varrho) = \{x : \|x - x_0\| \leq \varrho\}$

Theorem (Kantorovich, 1948, Heindl, Deuflhard, 1980)
Assume that in some open convex D

a) $\|f'(x_0)^{-1}\| \leq \beta$, $\|f'(x_0)^{-1} f(x_0)\| \leq \eta$	a) $\|f'(x_0)^{-1} f(x_0)\| \leq \eta$
b) $\|f'(u) - f'(v)\| \leq \kappa \cdot \|u - v\|$	b) $\|f'(x_0)^{-1}(f'(u) - f'(v))\| \leq \omega \cdot \|u - v\|$
c) $\omega = \beta\kappa,\ h = \eta\omega \leq \frac{1}{2}$	c) $h = \eta\omega \leq \frac{1}{2}$
(standard form)	(affine invariant form)

Newton-Kantorovich

$\mathcal{N}$ = norm ball, $\mathcal{N} = B(x_0, \varrho) = \{x : \|x - x_0\| \leq \varrho\}$

Theorem (Kantorovich, 1948, Heindl, Deuflhard, 1980)
Assume that in some open convex D

a) $\|f'(x_0)^{-1}\| \leq \beta$, $\|f'(x_0)^{-1} f(x_0)\| \leq \eta$	a) $\|f'(x_0)^{-1} f(x_0)\| \leq \eta$
b) $\|f'(u) - f'(v)\| \leq \kappa \cdot \|u - v\|$	b) $\|f'(x_0)^{-1} (f'(u) - f'(v))\| \leq \omega \cdot \|u - v\|$
c) $\omega = \beta\kappa,\ h = \eta\omega \leq \frac{1}{2}$	c) $h = \eta\omega \leq \frac{1}{2}$
(standard form)	(affine invariant form)

Let $\varrho = \left(1 - \sqrt{1 - 2h}\right) / \omega$.
If $B(x_0, \varrho) \subset D$, there exists a zero x^* of f in $B(x_0, \varrho)$.

Miranda's Theorem

$\mathcal{N} = [x]$ interval vector, $[x] = [\underline{x}, \overline{x}]$

$$[x]_i^+ = \{x \in [x] : x_i = \overline{x}_i\},\ i = 1, \ldots, n$$
$$[x]_i^- = \{x \in [x] : x_i = \underline{x}_i\},\ i = 1, \ldots, n$$

Miranda's Theorem

$\mathcal{N} = [x]$ interval vector, $[x] = [\underline{x}, \overline{x}]$

$$\begin{aligned} [x]_i^+ &= \{x \in [x] : x_i = \overline{x}_i\},\ i = 1, \ldots, n \\ [x]_i^- &= \{x \in [x] : x_i = \underline{x}_i\},\ i = 1, \ldots, n \end{aligned}$$

Theorem (Miranda, 1940)

f continuous on $[x]$. Assume that for $i = 1, \ldots, n$

$$f_i(x) \geq 0,\ f_i(y) \leq 0 \text{ for all } x \in [x]_i^+, y \in [x]_i^-$$

Then f has a zero x^* in $[x]$.

(n-dimensional version of the intermediate value theorem)

A First Hierarchy

Convention

$$\text{Theorem A} \quad \longrightarrow \text{Theorem B}$$

means

$$\text{Hypothesis of A} \Rightarrow \text{Hypothesis of B}$$

In other words:

A First Hierarchy

Convention

$$\text{Theorem A} \quad \longrightarrow \text{Theorem B}$$

means

$$\text{Hypothesis of A} \Rightarrow \text{Hypothesis of B}$$

In other words:
Theorem A is a special case of Theorem B

A First Hierarchy

Convention

$$\text{Theorem A} \quad \longrightarrow \text{Theorem B}$$

means

$$\text{Hypothesis of A} \Rightarrow \text{Hypothesis of B}$$

In other words:
Theorem A is a special case of Theorem B
Theorem B is more general

A First Hierarchy

Convention

Theorem A $\longrightarrow$ Theorem B

means

Hypothesis of A $\Rightarrow$ Hypothesis of B

In other words:
Theorem A is a special case of Theorem B
Theorem B is more general
Theorem B is a stronger theorem

A First Hierarchy

Convention

$$\text{Theorem A} \quad \longrightarrow \text{Theorem B}$$

means

$$\text{Hypothesis of A} \Rightarrow \text{Hypothesis of B}$$

In other words:
Theorem A is a special case of Theorem B
Theorem B is more general
Theorem B is a stronger theorem
Hypothesis of B is weaker

A First Hierarchy

Convention

$$\text{Theorem A} \qquad \longrightarrow \text{Theorem B}$$

means

$$\text{Hypothesis of A} \Rightarrow \text{Hypothesis of B}$$

In other words:
Theorem A is a special case of Theorem B
Theorem B is more general
Theorem B is a stronger theorem
Hypothesis of B is weaker

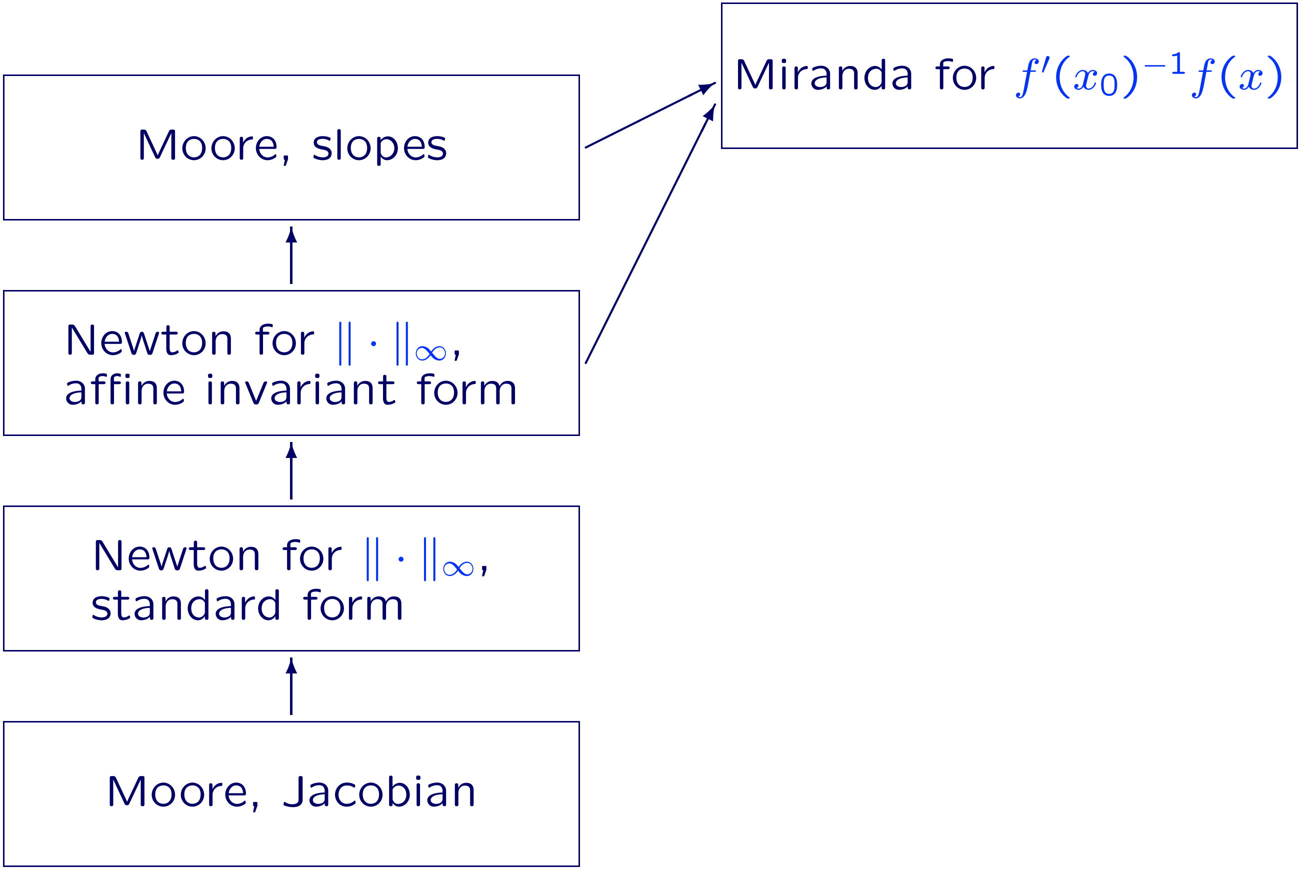
It's **better** to **be hit** by the arrow $\longrightarrow$

Moore, $[Y] = \Box\{f'(x), x \in [x]\}$ $\longrightarrow$ Newton-Kantorovich
for $\|\cdot\|_\infty$
standard form
(Rall, 1980)

Moore, $[Y] = \Box\{\delta f(x, x_0), x \in [x]\}$ $\longleftarrow$ Newton-Kantorovich
$(\delta f(x, x_0) = \int_0^1 f'(x^0 + t(x - x_0)\ dt)$ for $\|\cdot\|_\infty$
standard form
affine invariant form
(Neumaier and Shen, 1990, Shen and Wolfe, 1990)

Moore for f, slopes for $[Y]$ $\longrightarrow$ Miranda for Af
(Alefeld, Shen 2001)

Newton-Kantorovich for $\|\cdot\|_\infty$ $\longrightarrow$ Miranda for $f'(x_0)^{-1}f(x)$
standard or affine invariant
(Alefeld, Potra, Shen, 2001)


Miranda for $f'(x_0)^{-1} f(x)$
Moore, slopes
Newton for $\|\cdot\|_\infty$, affine invariant form
Newton for $\|\cdot\|_\infty$, standard form
Moore, Jacobian

Precise sample

Theorem: Let

$$\|f'(x_0)^{-1}f(x_0)\|_\infty \le \eta$$
$$\|f'(x_0)^{-1}(f'(u) - f'(v))\|_\infty \le \omega\|u - v\|_\infty$$
$$h = \eta\omega \le \frac{1}{2}$$
$$\varrho_- = (1 - \sqrt{1-2h})/\omega,\ \varrho_+ = (1 + \sqrt{1-2h})/\omega$$

Assume $\eta \neq 0$.
Take $\varrho \in [\varrho_-, \varrho_+]$, $\quad \mathcal{N} = [x] = x_0 + \varrho[-1,1]^n$.
Take $g(x) = f'(x^0)^{-1}f(x)$
Then for $i = 1, \dots, n$

$$g_i(x) \ge 0 \text{ for } x \in [x]_i^+$$
$$g_i(x) \le 0 \text{ for } y \in [x]_i^-$$

Question: Relate Newton-Kantorovich for arbitrary norm to “generalized Miranda“

Question: Relate Newton-Kantorovich for arbitrary norm to “generalized Miranda“

Dual norm: $\|\cdot\| \leftrightarrow \|\cdot\|_d$

$$\|y\|_d = \max_{\|x\|=1} |\langle x, y\rangle|$$

Generalized Miranda

Theorem

Let $f : B(x_0, \varrho) \to \mathbb{R}^N$ be continuous s.t. for all $x \in \partial B$

$$\langle a, f(x) \rangle \geq 0 \text{ for all } a \text{ with } \langle x - x_0, a \rangle = \|x - x_0\| \cdot \|a\|_d$$

Then f has a zero x^* in $B(x_0, \varrho)$.

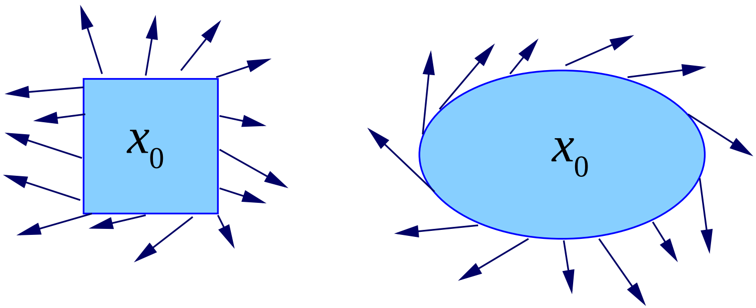
Generalized Miranda

Theorem

Let $f : B(x_0, \varrho) \to \mathbb{R}^N$ be continuous s.t. for all $x \in \partial B$

$$\langle a, f(x)\rangle \geq 0 \text{ for all } a \text{ with } \langle x - x_0, a\rangle = \|x - x_0\| \cdot \|a\|_d$$

Then f has a zero x^* in $B(x_0, \varrho)$.

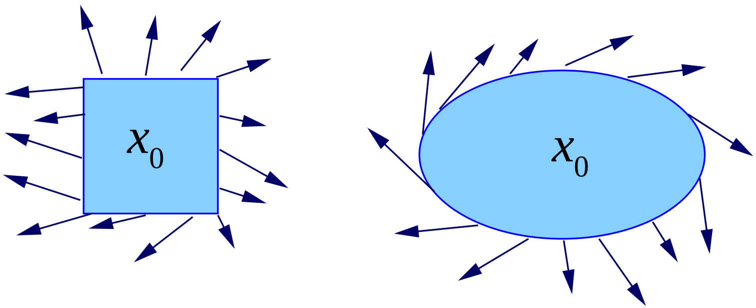


Generalized Miranda

Theorem
Let $f : B(x_0, \varrho) \to \mathbb{R}^N$ be continuous s.t. for all $x \in \partial B$

$$\langle a, f(x)\rangle \geq 0 \text{ for all } a \text{ with } \langle x - x_0, a\rangle = \|x - x_0\| \cdot \|a\|_d$$

Then f has a zero x^* in $B(x_0, \varrho)$.



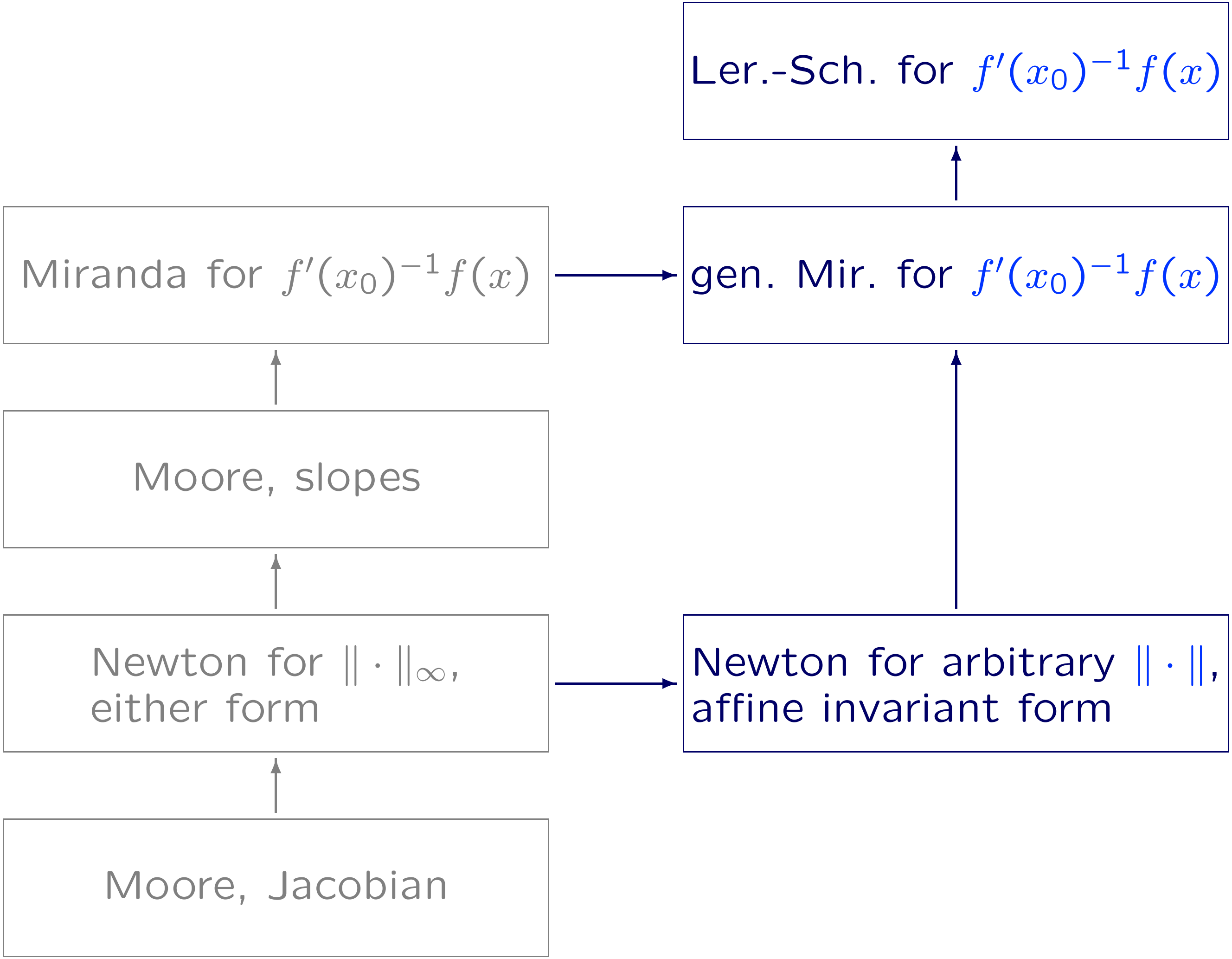
Theorem (Leray-Schauder) Let $\Omega \subseteq \mathbb{R}^n$ be open and bounded, $f : \overline{\Omega} \to \mathbb{R}^n$ continuous, $x_0 \in \Omega$.
Assume $f(x) = \lambda(x - x_0)$ for $x \in \partial\Omega \Rightarrow \lambda \geq 0$.
Then f has a zero x^* in Ω.

Theorem: Let

$$\|f'(x_0)^{-1}f(x_0)\| \le \eta$$
$$\|f'(x_0)^{-1}(f'(u)-f'(v))\| \le \omega\|u-v\|$$
$$h = \eta\omega \le \frac{1}{2}$$
$$\varrho_- = (1-\sqrt{1-2h})/\omega,\ \varrho_+ = (1+\sqrt{1-2h})/\omega$$

Assume $\eta \neq 0, \omega \neq 0$.
Take $\varrho \in [\varrho_-, \varrho_+], \quad \mathcal{N} = B(x_0, \varrho)$.
Take $g(x) = f'(x^0)^{-1}f(x)$
Then for all $x \in \partial B$

$\langle a, g(x)\rangle \ge 0$ for all a with $\langle x - x_0, a\rangle = \|x - x_0\| \cdot \|a\|_d$


Ler.-Sch. for $f'(x_0)^{-1}f(x)$
Miranda for $f'(x_0)^{-1}f(x)$
gen. Mir. for $f'(x_0)^{-1}f(x)$
Moore, slopes
Newton for $\|\cdot\|_\infty$, either form
Newton for arbitrary $\|\cdot\|$, affine invariant form
Moore, Jacobian

Affine invariance

Observation:
Generalized Miranda, Leray-Schauder are not affine invariant but we apply theorems to $f'(x_0)^{-1}f(x)$, which is affine invariant

Question: Are there other affine invariant theorems around
$\longrightarrow$ Borsuk's theorem

Theorem (Borsuk, 1933): Let $\Omega \subseteq \mathbb{R}^n$ open, convex and symmetric w.r.t. $x_0 \in \Omega$, $f : \overline{\Omega} \to \mathbb{R}^n$ continuous. Assume for all $x_0 + y \in \partial\Omega$

$$f(x_0 + y) \neq \lambda \cdot f(x_0 - y) \text{ with } \lambda \in [1, \infty)$$

Then f has a zero x^* in $\overline{\Omega}$.

Note: $f(x_0+y) \neq \lambda \cdot f(x_0-y) \leftrightarrow Af(x_0+y) \neq \lambda Af(x_0+y)$

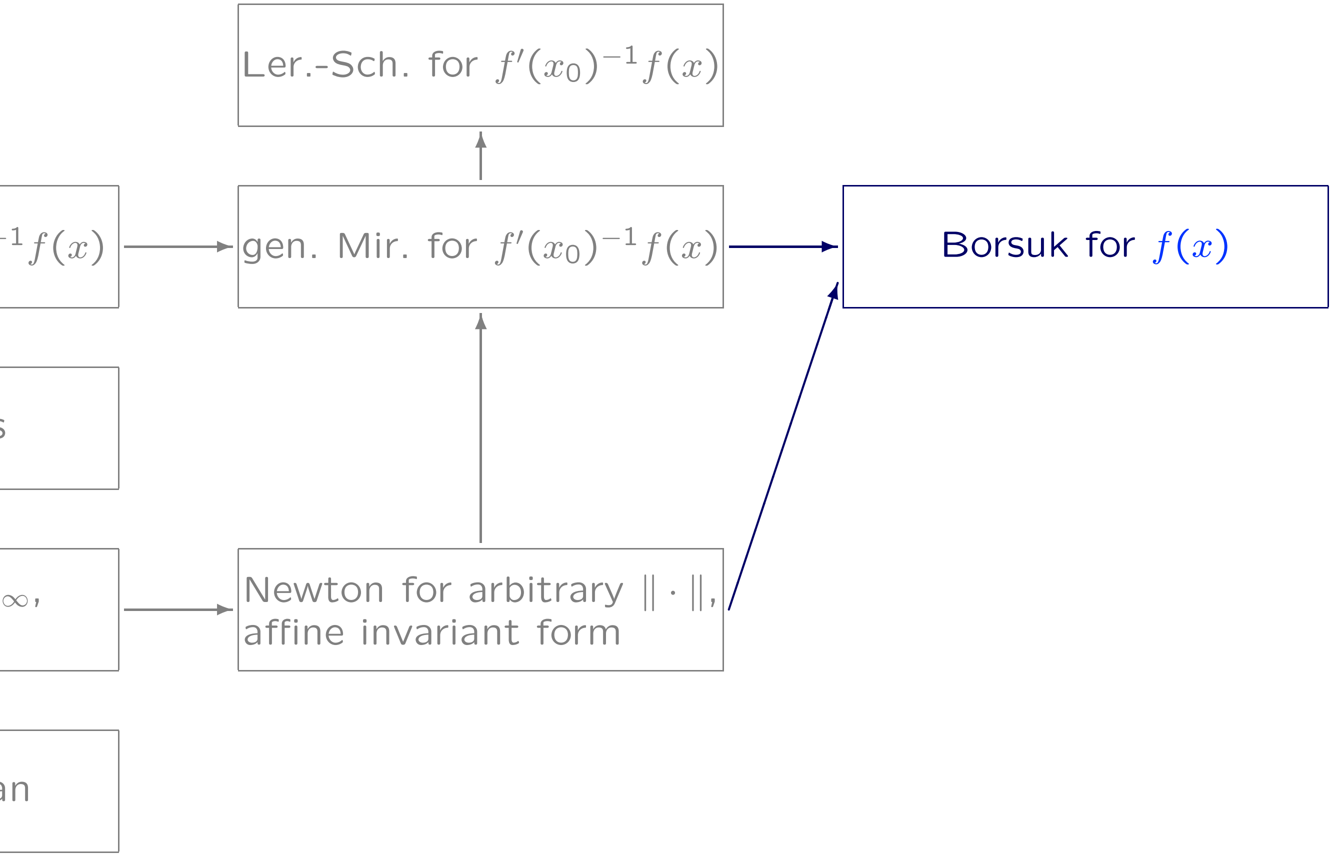
Borsuk's theorem is affine invariant.

Theorem: Let

$$\|f'(x_0)^{-1}f(x_0)\| \leq \eta$$
$$\|f'(x_0)^{-1}(f'(u) - f'(v))\| \leq \omega\|u - v\|$$
$$h = \eta\omega \leq \frac{1}{2}$$
$$\varrho_- = (1 - \sqrt{1 - 2h})/\omega,\ \varrho_+ = (1 + \sqrt{1 - 2h})/\omega$$

Then for all $x_0 + y \in \partial B(x_0, \varrho)$

$$f(x_0 + y) \neq \lambda \cdot f(x_0 - y) \text{ with } \lambda \in [1, \infty)$$


Ler.-Sch. for $f'(x_0)^{-1}f(x)$
gen. Mir. for $f'(x_0)^{-1}f(x)$
Borsuk for $f(x)$
Newton for arbitrary $\|\cdot\|$,
affine invariant form

Practical Aspects

Miranda's theorem $(\mathcal{N} = [x])$

Take $g(x) = f'(x_0)^{-1} f(x)$

Mean value extension

$$\{g(y) : y \in [y] \subseteq [x]\} \subseteq g(x_0) + g'([x])([y] - x_0) =: mg([y])$$

Test:

for $i = 1, \ldots, n$

check for $mg_i([x]_i^+) \geq 0$, $mg_i([x]_i^-) \leq 0$

(Moore, Kioustelidis, 1980)

Remember: $mg_i([x]_i^+) \geq 0 \Leftrightarrow g_i(x_0) + g_i'([x])([x]_i^+ - x_0) \geq 0$.

New idea: Check for Borsuk on $[x]$

$$\begin{aligned} g_j(x_0+y) &= \lambda g_j(x_0-y) \\ \Leftrightarrow\ g_j(x_0)+g_j'(\xi)(y) &= \lambda g_j(x_0) - \lambda g_j'(\eta) y \\ \Leftrightarrow\quad (\lambda-1)\cdot g_j(x_0) &= \left(\lambda g_j'(\eta)+g_j'(\xi)\right) y \\ \Rightarrow\quad (\lambda-1)\cdot g_j(x_0) &\in \left(\lambda g_j'([x])+g_j'([x])\right) y \\ \Leftrightarrow\quad (\lambda-1)\cdot g_j(x_0) &\in (1+\lambda)\cdot g_j'([x]) y \\ \Rightarrow\quad \tfrac{\lambda-1}{\lambda+1}\cdot g_j(x_0) &\in g_j'([x])([x]_i^{\pm}-x_0) \end{aligned}$$

Note: $\frac{\lambda-1}{\lambda+1} \in (-1,1)$ for $\lambda > 0$.

New idea: Check for Borsuk on $[x]$

$$
\begin{aligned}
g_j(x_0+y) &= \lambda g_j(x_0-y) \\
\Leftrightarrow\ g_j(x_0)+g_j'(\xi)(y) &= \lambda g_j(x_0)-\lambda g_j'(\eta)y \\
\Leftrightarrow\ (\lambda-1)\cdot g_j(x_0) &= \big(\lambda g_j'(\eta)+g_j'(\xi)\big)\,y \\
\Rightarrow\ (\lambda-1)\cdot g_j(x_0) &\in \big(\lambda g_j'([x])+g_j'([x])\big)\,y \\
\Leftrightarrow\ (\lambda-1)\cdot g_j(x_0) &\in (1+\lambda)\cdot g_j'([x])y \\
\Rightarrow\ \tfrac{\lambda-1}{\lambda+1}\cdot g_j(x_0) &\in g_j'([x])([x]_i^{\pm}-x_0)
\end{aligned}
$$

Note: $\frac{\lambda-1}{\lambda+1}\in(-1,1)$ for $\lambda>0$.

Test: For $i=1,\ldots,n$ check for

$$\bigcap_{j=1}^{n}\frac{g_j'([x])([x]_i^{+}-x_0)}{g_j(x_0)}\cap(-1,1)=\emptyset$$

or

$$\bigcap_{j=1}^{n}\frac{g_j'([x])([x]_i^{-}-x_0)}{g_j(x_0)}\cap(-1,1)=\emptyset$$

Simplified version

Test for

$$\frac{g_i'([x])([x]_i^+ - x_0)}{g_i(x_0)} \cap (-1,1) = \emptyset \qquad \text{or}$$

$$\frac{g_i'([x])([x]_i^- - x_0)}{g_i(x_0)} \cap (-1,1) = \emptyset$$

which is equivalent to (case $g_i(x^0) > 0$)

$$g_i(x_0) + g_i'([x]) \cdot ([x]_i^+ - x_0) \leq 0 \qquad \text{or}$$
$$g_i(x_0) - g_i'([x]) \cdot ([x]_i^+ - x_0) \leq 0 \qquad \text{or}$$
$$g_i(x_0) + g_i'([x]) \cdot ([x]_i^- - x_0) \leq 0 \qquad \text{or}$$
$$g_i(x_0) - g_i'([x]) \cdot ([x]_i^- - x_0) \leq 0$$

Simplified version

Test for

$$\frac{g_i'([x])([x]_i^+ - x_0)}{g_i(x_0)} \cap (-1,1) = \emptyset \qquad \text{or}$$

$$\frac{g_i'([x])([x]_i^- - x_0)}{g_i(x_0)} \cap (-1,1) = \emptyset$$

which is equivalent to (case $g_i(x^0) > 0$)

$$\begin{aligned} g_i(x_0) + g_i'([x]) \cdot ([x]_i^+ - x_0) &\leq 0 \qquad \text{or} \\ g_i(x_0) - g_i'([x]) \cdot ([x]_i^+ - x_0) &\leq 0 \qquad \text{or} \\ g_i(x_0) + g_i'([x]) \cdot ([x]_i^- - x_0) &\leq 0 \qquad \text{or} \\ g_i(x_0) - g_i'([x]) \cdot ([x]_i^- - x_0) &\leq 0 \end{aligned}$$

Miranda test:

$$\begin{aligned} g_i(x_0) + g_i'([x]) \cdot ([x]_i^+ - x_0) &\geq 0 \qquad \text{and} \\ g_i(x_0) + g_i'([x]) \cdot ([x]_i^- - x_0) &\leq 0 \end{aligned}$$

Consequence: Borsuk test is better than Miranda test!

Conclusions

Conclusions

- Newton-Kantorovich $\longrightarrow$ Miranda in any norm

Conclusions

- Newton-Kantorovich $\longrightarrow$ Miranda in any norm
- Borsuk and Leray-Schauder are even more general

Conclusions

- Newton-Kantorovich $\longrightarrow$ Miranda in any norm
- Borsuk and Leray-Schauder are even more general
- Borsuk-based existence test using interval arithmetic improves upon Miranda

Conclusions

- Newton-Kantorovich $\longrightarrow$ Miranda in any norm
- Borsuk and Leray-Schauder are even more general
- Borsuk-based existence test using interval arithmetic improves upon Miranda
- Future work: Hierarchy of existence tests?