TECHNISCHE UNIVERSITÄT DRESDEN

Institute of Numerical Mathematics

An Extended Local Convergence Theory for Newton-type Methods

Andreas Fischer

joint work with Nico Strasdat

10th Workshop Dresden–Prague on Numerical Analysis

Děčín, November 4–5, 2022

Contents

Constrained Equation

$$t(z) = 0 \quad \text{subject to} \quad z \in \Omega$$

- $t : \mathbb{R}^N \to [0, \infty)$ is a given continuous mapping
- $\Omega \subseteq \mathbb{R}^N$ is a given nonempty closed set
- $Z := \{z \in \Omega \mid t(z) = 0\}$ denotes the solution set

General Iteration Scheme

Let the sequence $\{z^k\} \subset \Omega$ be generated by

$$z^{k+1} := z^+(z^k) \quad \text{for } k = 0, 1, 2, \ldots$$

- $z^+ : \Omega \to \Omega$ is a given mapping
- $z^0 \in \Omega$ is sufficiently close to some $z^* \in Z$

What assumptions on the mappings $t : \mathbb{R}^N \to [0, \infty)$ and $z^+ : \Omega \to \Omega$ yield fast local convergence of the sequence $\{z^k\}$ to the solution set Z ?

Local Convergence

Theorem

Suppose that $C \geq 1$, $\delta > 0$, and $\sigma > 1$ exist so that the conditions

$$\begin{array}{lrcl} \text{(a)} & \|z^+(z) - z\| & \leq & C \cdot t(z), \\ \text{(b)} & t(z^+(z)) & \leq & C \cdot t(z)^\sigma \end{array}$$

are satisfied for all $z \in \mathcal{B}(z^*, \delta) \cap \Omega$. Then, there exists $\varepsilon > 0$ so that, for any $z^0 \in \mathcal{B}(z^*, \varepsilon) \cap \Omega$,

- $\{z^k\}$ belongs to $\mathcal{B}(z^*, \delta) \cap \Omega$ and
- $\{z^k\}$ converges to some $\hat{z} \in Z$ with an R-order of at least σ.

Suppose, in addition, that t is Hölder continuous on $\mathcal{B}(z^*, \delta)$, i.e., $L_0 > 0$ and $\theta \in (0, 1]$ exist so that

$$|t(z) - t(z')| \leq L_0 \|z - z'\|^\theta \quad \text{for all } z, z' \in \mathcal{B}(z^*, \delta). \tag{1}$$

Then, if $\theta \cdot \sigma^\nu > 1$,

- $\{z^k\}$ converges to $\hat{z}$ with ν-step Q-order of at least $\theta \cdot \sigma^\nu$.

TECHNISCHE UNIVERSITÄT DRESDEN

Complementarity System

$$F(x) = 0, \quad a(x) \geq 0, \quad b(x) \geq 0, \quad a(x)^\top b(x) = 0$$

$F : \mathbb{R}^n \to \mathbb{R}^m$ and $a, b : \mathbb{R}^n \to \mathbb{R}^p$ are assumed to be differentiable with locally Lipschitz continuous Jacobians ($C^{1,1}$)

Complementarity Function (C-Function)

$$\varphi : \mathbb{R}^2 \to \mathbb{R} \quad \text{so that} \quad \varphi(c,d) = 0 \quad \Leftrightarrow \quad c \geq 0,\ d \geq 0,\ c \cdot d = 0$$

Reformulation of a Complementarity System

$$T_\varphi(z) := \begin{pmatrix} H(z) \\ \Phi_\varphi(u,v) \end{pmatrix} = 0, \qquad z = (x,u,v) \in \Omega := \underline{\mathbb{R}^n \times \mathbb{R}^p_+ \times \mathbb{R}^p_+}$$

with

$$H(z) := \begin{pmatrix} F(x) \\ a(x) - u \\ b(x) - v \end{pmatrix}, \qquad \Phi_\varphi(u,v) := \begin{pmatrix} \varphi(u_1, v_1) \\ \vdots \\ \varphi(u_p, v_p) \end{pmatrix} = 0$$

A solution $z^* = (x^*, u^*, v^*)$ is called **degenerate** if $u_i^* = v_i^* = 0$ for at least one index $i \in \{1, \dots, p\}$

Examples of C-Functions

min-function (Pang 1991)

$$\varphi_{\min}(a,b) := \min\{a,b\}$$

$\varphi_{\min}$ is piecewise smooth

Fischer-Burmeister function (1992)

$$\varphi_{FB}(a,b) := \sqrt{a^2+b^2} - a - b$$

φ_{FB} is **not** piecewise smooth

- If z^* is a degenerate solution, a smooth C-function φ implies that T_φ has a singular Jacobian at z^* and $\|T_\varphi\|$ cannot provide an error bound around z^*.
- Many other C-functions exist (Galantai 12), for example

Mangasarian 76	$\varphi(a,b) := \lvert a-b\rvert - b - a = -2\varphi_{\min}(a,b)$
Evtushenko&Purtov 84	$\varphi(a,b) := -ab + \frac{1}{2}\min\{0,a+b\}^2$
Mangasarian&Solodov 93	$\varphi(a,b) := ab - a^2 - b^2 + \frac{1}{2\alpha}\left(\max\{0,a-\alpha b\}^2 + \max\{0,b-\alpha a\}^2\right),\ \alpha > 1$
Luo&Tseng 96	$\|(a,b)\|_p - a - b,\ p > 1$
Chen&Chen&Kanzow 97	$\varphi(a,b) := \varphi_{FB} - \beta\max\{0,a\}\max\{0,b\},\ \beta > 0$

Difficulties for Newton-type Methods

- a degenerate solution
- a nonisolated solution
- a solution which is degenerate and nonisolated

Sources of Nonisolatedness

- Redundancies in decision variables
- Nonuniqueness of Lagrange mulipliers
- Generic nonisolatedness of solutions like in generalized Nash Equilibrium Problems

Older Results for the Unconstrained Case ($\Omega = \mathbb{R}^n$)

$\varphi \in \{\varphi_{\min}, \varphi_{FB}\}$

- **Newton's method**

$$T_\varphi(z^k) + T'_\varphi(z^k)(z - z^k) = 0$$

  quadratic convergence if z^* is not degenerate and $T'_\varphi(z^*)$ is nonsingular

- **Semismooth Newton method** (Kummer 1988, Qi/Sun 1993, ...)

$$T_\varphi(z^k) + V_k(z - z^k) = 0 \quad \text{with} \quad V_k \in \partial T_\varphi(z^k)$$

  quadratic convergence if $\partial T_\varphi(z^*)$ is nonsingular

- **Levenberg-Marquardt method** (Yamashita/Fukushima 2001)

$$\min_z \|T_\varphi(z^k) + T'_\varphi(z^k)(z - z^k)\|^2 + \lambda(z^k)\|z - z^k\|^2 \quad \text{with } \lambda(z^k) := \|T_\varphi(z^k)\|^2$$

  quadratic convergence if z^* is not degenerate and T_φ provides an **Error Bound**, i.e., there are $\delta > 0, \omega > 0$ so that

$$\|T_\varphi(z)\| \geq \omega \operatorname{dist}[z, Z] \quad \text{for all } z \in \mathcal{B}(z^*, \delta)$$

Methods for the Constrained Case

Constrained Levenberg-Marquardt method (Kanzow/Yamashita/Fukushima 2004)

$$\min_{z\in\Omega} \|T_\varphi(z^k) + G_\varphi(z^k)(z - z^k)\|^2 + \lambda(z^k)\|z - z^k\|^2$$

with $\lambda(z^k) := \|T_\varphi(z^k)\|^2$

LP-Newton method (Facchinei/F./Herrich 2014)

$$\min_{z,\gamma} \mu \quad \text{s.t.} \quad \begin{array}{l} \|T_\varphi(z^k) + G_\varphi(z^k)(z - z^k)\|_\infty \leq \mu \|T_\varphi(z^k)\|_\infty^2 \\ \|z - z^k\|_\infty \leq \mu \|T_\varphi(z^k)\|_\infty \\ z \in \Omega,\ \mu \geq 0 \end{array}$$

- G_φ Jacobian or substitute of Jacobian of T_φ
- Subproblems are strongly convex quadratic programs or linear programs. They always have a solution.

Local Convergence for the Constrained Case

Assumptions

- $\boxed{\varphi = \varphi_{\min}}$, i.e. T_φ is a piecewise $C^{1,1}$ mapping
- $G_\varphi(z) \in \{(T_\varphi^i)'(z) \mid T_\varphi^i(z) = T_\varphi(z)\}$
- **Piecewise Constrained Error Bound**:

  There are $\delta, \omega > 0$ so that

  $$\omega \mathrm{dist}[z, Z_i] \le \|T_\varphi^i(z)\| \quad \text{for all } z \in \mathcal{B}(z^*, \delta) \cap \Omega$$

  holds for all i with $z^* \in Z_i := \{z \in \Omega \mid T_\varphi^i(z) = 0\}$ (Z_i = **solution set of a "piece").**

Convergence

There is $\varepsilon > 0$ so that, for any $z^0 \in \mathcal{B}(z^*, \varepsilon)$, the LP-Newton method and the constrained Levenberg–Marquardt method are well-defined and generate a sequence $\{z^k\}$ that converges Q-quadratically to some $\hat{z} \in Z$.

Assumptions allow cases with solutions that are degenerate and nonisolated.

(Facchinei/F./Herrich 2013+14, F./Herrich/Izmailov/Solodov 2016)

TECHNISCHE UNIVERSITÄT DRESDEN

Specification of the Algorihmic Framework

- $\boxed{\varphi := \varphi_{FB}}$
- For any $z \in \mathbb{R}^N$, the mapping $t : \mathbb{R}^N \to [0, \infty)$ is defined by

$$t(z) := \| T_\varphi(z)\|^{2/3}.$$

- The mapping $z^+ : \Omega \to \Omega$ is defined as $z^+(z) := z$ for $z \in Z$ and, for $z \in \Omega \setminus Z$, as the unique solution of the constrained Levenberg-Marquardt subproblem

$$\min_{\xi \in \Omega} \| T_\varphi(z) + G_\varphi(\xi - z)\|^2 + \lambda(z)\|\xi - z\|^2$$

with $\lambda(z) := \| T(z)\|^\gamma$ for $\gamma := 4/3$ and $G_\varphi(z) \in \partial T_\varphi(z)$.

Index Error Bound

Let $z^* = (x^*, u^*, v^*) \in Z$ be a fixed solution. For index sets $I, J \subseteq \{1, \ldots, p\}$, let us define

$$T_{IJ}(z) := \begin{pmatrix} H(z) \\ u_I \\ v_J \end{pmatrix} \quad \text{and} \quad Z_{IJ} := \{z \in \Omega \mid T_{IJ}(z) = 0\}$$

Assumption (Index Error Bound)

There exist $\omega > 0$ and $\delta > 0$ so that, for all (I, J) with $I \cup J = \{1, \ldots, p\}$ and $I \subseteq \{i \mid u_i^* = 0\}$, $J \subseteq \{j \mid v_j^* = 0\}$, it holds that

$$\omega \operatorname{dist}[z, Z_{IJ}] \leq \| T_{IJ}(z) \| \qquad \text{for all } z \in \mathcal{B}(z^*, \delta) \cap \Omega.$$

The Index Error Bound implies the Piecewise Constrained Error Bound, but still allows cases with solutions that are degenerate and nonisolated.

Convergence Result from the Algorithmic Framework

Theorem

Let the Index Error Bound be satisfied.
Then, there exists $\varepsilon > 0$ such that, for any $z^0 \in \mathcal{B}(z^*, \varepsilon) \cap \Omega$, the sequence $\{z^k\}$ generated by the constrained Levenberg–Marquardt method for $\lambda(z) := \|T_\varphi(z)\|^{4/3}$ is well-defined and converges to some $\hat{z} \in Z$ with

- an R-order of $4/3$,
- a 4-step Q-order of $2.1\ldots$,
- $\{\|T_\varphi(z^k)\|\}$ and $\{\text{dist}[z^k, Z]\}$ converge to 0 with a Q-order of $4/3$.

TECHNISCHE UNIVERSITÄT DRESDEN

Final Remarks

- The Algorithmic Framework also covers several existing results on local convergence results for complementarity systems.
- In contrast to usual convergence theories for algorithms with superlinear convergence, the Algorithmic Framework allows **long** steps, i.e.,

$$\frac{\|z^{k+1} - z^k\|}{\mathrm{dist}[z^k, Z]}$$

  can be become unbounded for $k \to \infty$.
- A result in Behling/F./Schönefeld/Strasdat (2019) with long steps for the multipliers only is significantly generalized.
- It is difficult to find an example for which the constrained Levenberg–Marquardt method does not converge with a local quadratic rate.

TECHNISCHE UNIVERSITÄT DRESDEN

Some Related References

Fischer, A., Strasdat, N.: An extended convergence framework applied to complementarity systems with degenerate and nonisolated solutions. Pure and Applied Functional Analysis, to appear

Fischer, A., Herrich, M., Izmailov, A.F., Solodov, M.V.: Convergence conditions for Newton-type methods applied to complementarity systems with nonisolated solutions. Comput. Optim. Appl. 63 (2016) 425–459.

Behling, R., Fischer, A., Schönefeld, K., Strasdat, N.: A special complementarity function revisited. Optimization 68 (2019) 65–79.

Facchinei, F., Fischer, A., Herrich. M.: An LP–Newton method: Nonsmooth equations, KKT systems, and nonisolated solutions. Math. Program. 146 (2014) 1–36.

Facchinei, F., Fischer, A., Herrich. M.: A family of Newton methods for nonsmooth constrained systems with nonisolated solutions. Math. Methods Oper. Res. 77 (2013) 433–443.