Towards a general convergence theory for inexact Newton regularizations

Andreas Rieder

Institut für Angewandte und Numerische Mathematik

Universität Karlsruhe
Fakultät für Mathematik

(jointly with Armin Lechleiter, Palaiseau)

Overview

REGINN: An inexact Newton regularization

Level set based termination

Local convergence

Bibliographical notes

Conclusion

REGINN: An inexact Newton regularization

Newton regularizations

$$F : \mathsf{D}(F) \subset X \to Y, \quad X, Y \text{ Hilbert spaces}$$

$$F(x) = y^\delta$$

where $\|y - y^\delta\|_Y \leq \delta$, $y = F(x^+)$, and $F(x) = y$ locally ill-posed in x^+.

Let x_n be an approximation to x^+: $\quad x_{n+1} = x_n + s_n^{\mathrm{N}}$

The exact Newton step $s_n^{\mathsf{e}} = x^+ - x_n$ satisfies $\left(A_n := F'(x_n) \right)$

$$A_n s_n^{\mathsf{e}} = y - F(x_n) - E(x^+, x_n)$$

Newton regularizations

$$F : \mathrm{D}(F) \subset X \to Y, \quad X, Y \text{ Hilbert spaces}$$

$$F(x) = y^\delta$$

where $\|y - y^\delta\|_Y \leq \delta$, $y = F(x^+)$, and $F(x) = y$ locally ill-posed in x^+.

Let x_n be an approximation to x^+: $\quad x_{n+1} = x_n + s_n^{\mathrm{N}}$

The exact Newton step $s_n^{\mathsf{e}} = x^+ - x_n$ satisfies $\left(A_n := F'(x_n) \right)$

$$A_n s_n^{\mathsf{e}} = y - F(x_n) - E(x^+, x_n)$$

$\Longrightarrow$ Determine s_n^{N} as regularized solution of

$$A_n s = b_n^\delta, \quad b_n^\delta := y^\delta - F(x_n)$$

Newton regularizations

$$F : \mathrm{D}(F) \subset X \to Y, \quad X, Y \text{ Hilbert spaces}$$

$$F(x) = y^\delta$$

where $\|y - y^\delta\|_Y \le \delta$, $y = F(x^+)$, and $F(x) = y$ locally ill-posed in x^+.

Let x_n be an approximation to x^+: $\quad x_{n+1} = x_n + s_n^{\mathrm{N}}$

The exact Newton step $s_n^{\mathsf{e}} = x^+ - x_n$ satisfies $(A_n := F'(x_n))$

$$A_n s_n^{\mathsf{e}} = y - F(x_n) - E(x^+, x_n)$$

$\Longrightarrow$ Determine s_n^{N} as regularized solution of

$$A_n s = b_n^\delta, \quad b_n^\delta := y^\delta - F(x_n)$$

Let $\{s_{n,m}\}_{m \in \mathbb{N}}$ a regularizing sequence. Then, $s_n^{\mathrm{N}} = s_{n,m_n}$.

For instance, $s_{n,m} = g_m(A_n^* A_n) A_n^* b_n^\delta$ where $g_m : [0, \|A_n\|^2] \to \mathbb{R}$ is a so-called filter function.

$$
\begin{array}{l}
\texttt{REGINN}(x_{N(\delta)}, R, \{\mu_n\}) \\
n := 0; \;\; x_0 := x_{N(\delta)}; \\
\texttt{while} \;\; \|b_n^\delta\|_Y > R\delta \;\; \texttt{do} \\
\{ \quad m := 0, \;\; s_{n,0} = 0; \\
\quad \texttt{repeat} \\
\qquad m := m + 1; \\
\qquad \texttt{compute} \;\; s_{n,m} \;\; \texttt{from} \;\; A_n s = b_n^\delta; \\
\quad \texttt{until} \;\; \|A_n s_{n,m} - b_n^\delta\|_Y < \mu_n \|b_n^\delta\|_Y \\
\quad x_{n+1} := x_n + s_{n,m}; \\
\quad n := n + 1; \\
\} \\
x_{N(\delta)} := x_n;
\end{array}
$$

$$
m_n = \min\left\{ m \in \mathbb{N} \,:\, \|A_n s_{n,m} - b_n^\delta\|_Y < \mu_n \|b_n^\delta\|_Y \right\}
$$

Assumptions on $\{s_{n,m}\}$

For the analysis of `REGINN` we require three properties of the regularizing sequence $\{s_{n,m}\}$, namely

1. $\langle A_n s_{n,m}, b_n^\delta \rangle_Y > 0 \;\; \forall m \geq 1$ whenever $A_n^* b_n^\delta \neq 0$,
2. $\lim_{m\to\infty} A_n s_{n,m} = P_{\overline{\mathsf{R}(A_n)}} b_n^\delta$,
3. $\exists\, \Theta \geq 1$: $\|A_n s_{n,m}\|_Y \leq \Theta \|b_n^\delta\|_Y \;\; \forall m, n$.

If $s_{n,m} = g_m(A_n^* A_n) A_n^* b_n^\delta$ and

$$0 < \lambda g_m(\lambda) \leq C_g,\; \lambda > 0, \quad \text{and} \quad \lim_{m\to\infty} g_m(\lambda) = 1/\lambda,\; \lambda > 0,$$

then all three requirements are fulfilled where $\Theta \leq C_g$.

Examples: Landweber, implicit iteration, Tikhonov, Showalter, ν-methods, as well as non-linear methods: steepest decent and conjugate gradients

Lemma: Any direction $s_{n,m}$ is a descent direction in x_n for the functional

$$\varphi(\cdot) = \|y^\delta - F(\cdot)\|_Y^2,$$

that is,

$$\langle \nabla\varphi(x_n), s_{n,m}\rangle_X < 0 \text{ for } m \geq 1 \text{ whenever } A_n^* b_n^\delta \neq 0.$$

Lemma: Assume that $\|P_{\mathsf{R}(A_n)^\perp} b_n^\delta\|_Y < \|b_n^\delta\|_Y$. Then, for any tolerance

$$\mu_n \in \left] \frac{\|P_{\mathsf{R}(A_n)^\perp} b_n^\delta\|_Y}{\|b_n^\delta\|_Y}, 1 \right[$$

the `repeat`-loop of `REGINN` terminates.

Remark: Under $\|P_{\mathsf{R}(A_n)^\perp} b_n^\delta\|_Y = \|b_n^\delta\|_Y$, that is, $\|P_{\overline{\mathsf{R}(A_n)}} b_n^\delta\|_Y = 0$ we have $s_{n,m} = 0$ for all m.

Level set based termination

Structural assumptions on non-linearity

For $x_0 \in \mathsf{D}(F)$ such that $\|F(x_0) - y^\delta\|_Y > \delta$ define the level set

$$\mathfrak{L}(x_0) := \big\{x \in \mathsf{D}(F) \colon \|F(x) - y^\delta\|_Y \le \|F(x_0) - y^\delta\|\big\}.$$

Note that $x^+ \in \mathfrak{L}(x_0)$.

Assume

$$\|F(v) - F(w) - F'(w)(v-w)\|_Y \le L\,\|F'(w)(v-w)\|_Y$$
$$\text{for one } L < 1 \text{ and for all } v, w \in \mathfrak{L}(x_0) \text{ with } v - w \in \mathsf{N}\big(F'(w)\big)^\perp$$

and

$$\big\|P_{\mathsf{R}(F'(u))^\perp}\big(y - F(u)\big)\big\|_Y \le \varrho\|y - F(u)\|_Y$$
$$\text{for one } \varrho < 1 \text{ and all } u \in \mathfrak{L}(x_0).$$

Remark:

$$L < \frac{1}{2} \quad\Longrightarrow\quad \varrho \le \frac{L}{1-L} < 1$$

Example

Let $\{v_n\}$ and $\{u_n\}$ be ONB in separable Hilbert spaces X and Y, resp.
We define operator $F\colon X \to Y$ by

$$F(x) = \sum_{n=1}^{\infty} \frac{1}{n} f\big(\langle x, v_n\rangle_X\big) u_n$$

where $f\colon \mathbb{R} \to \mathbb{R}$ is smooth with $f'(\cdot) \geq f'_{\min} > 0$.

Here, $\overline{\mathsf{R}(F'(x))} = Y$ for any $x \in X$. Thus, $\varrho = 0$.

If, further, $f'(\cdot) \leq f'_{\max}$ with $f'_{\max} < 2 f'_{\min}$ then $L = \frac{f'_{\max} - f'_{\min}}{f'_{\min}} < 1$.

Termination

Theorem: Let $\Theta L + \varrho < \Lambda$ for one $\Lambda < 1$. Further, choose

$$R > \frac{1+\varrho}{\Lambda - \Theta L - \varrho}.$$

Finally, select all tolerances $\{\mu_n\}$ such that

$$\mu_n \in \big]\mu_{\min,n},\ \Lambda - \Theta L\big], \quad \text{with } \mu_{\min,n} := \frac{(1+\varrho)\delta}{\|b_n^\delta\|_Y} + \varrho.$$

Then, there exists an $N(\delta)$ such that $\{x_1, \ldots, x_{N(\delta)}\} \subset \mathfrak{L}(x_0)$. Moreover, only the final iterate satisfies the discrepancy principle, that is,

$$\|y^\delta - F(x_{N(\delta)})\|_Y \le R\delta,$$

and

$$\frac{\|y^\delta - F(x_{n+1})\|_Y}{\|y^\delta - F(x_n)\|_Y} < \mu_n + \theta_n L \le \Lambda, \quad n = 0, \ldots, N(\delta) - 1,$$

where $\theta_n = \|A_n s_n^{\mathrm{N}}\|_Y / \|b_n^\delta\|_Y \le \Theta$.

Remark: Although $\|y - F(x_{N(\delta)})\|_Y < (R+1)\delta$ we do not have convergence of $\{x_{N(\delta)}\}$ as $\delta \to 0$ in general.

Local convergence

Additional assumptions on $\{s_{n,m}\}$

Monotonicity: Let there be a continuous and monotonically increasing function $\Psi\colon \mathbb{R} \to \mathbb{R}$ with $t \le \Psi(t)$ for $t \in [0,1]$ such that if $\gamma_n = \|b_n^\delta - A_n s_n^{\mathsf{e}}\|_Y / \|b_n^\delta\|_Y < 1$ and $\|b_n^\delta - A_n s_{n,m-1}\|_Y / \|b_n^\delta\|_Y \ge \Psi(\gamma_n)$ then

$$\|s_{n,m} - s_n^{\mathsf{e}}\|_X < \|s_{n,m-1} - s_n^{\mathsf{e}}\|_X.$$

Stability: $\lim\limits_{\delta \to 0} s_{n,m}(y^\delta) = s_{n,m}(y)$.

Examples:

- Landweber iteration and steepest decent: $\Psi(t) = 2t$,
- Implicit iteration: $\Psi(t) = Ct$ where $C > 2$,
- cg-method: $\Psi(t) = \sqrt{2t}$.

Modified structural assumption

Assume

$$\|F(v) - F(w) - F'(w)(v - w)\|_Y \leq L\|F'(w)(v - w)\|_Y$$

for one $L < 1$ and for all $v, w \in B_r(x^+) \subset \mathsf{D}(F)$.

Monotonicity and Convergence

Theorem: Let

$$\Psi\Big(\frac{L}{1-L}\Big) + \Theta L < \Lambda \quad \text{for one } \Lambda < 1$$

and define

$$\mu_{\min} := \Psi\bigg(\Big(\frac{1}{R} + L\Big)\frac{1}{1-L}\bigg).$$

Choose R so large that

$$\mu_{\min} + \Theta L < \Lambda.$$

Restrict all tolerances $\{\mu_n\}$ to $]\mu_{\min}, \Lambda - \Theta L]$ and start with $x_0 \in B_r(x^+)$. Then,

$$\|x^+ - x_n\|_X < \|x^+ - x_{n-1}\|_X, \quad n = 1, \ldots, N(\delta),$$

and, if x^+ is unique in $B_r(x^+)$,

$$\lim_{\delta \to 0} \|x^+ - x_{N(\delta)}\|_X = 0.$$

Bibliographical notes

Bibliographical notes

- B. Kaltenbacher, A. Neubauer, O. Scherzer
  Iterative Regularization Methods for Nonlinear Ill-posed Problems
  de Gruyter, Berlin, 2007
- A. Lechleiter, A. Rieder
  Towards a general convergence theory for inexact Newton regularizations
  Numer. Math., to appear (`download from my webpage`),
- M. Hanke
  Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems
  Numer. Funct. Anal. Optim. 18 (1998), 971-993.
- Q. Jin and U. Tautenhahn
  On the discrepancy principle for some Newton type methods for solving nonlinear ill-posed problems
  Numer. Math., 111 (2009), 509-558.

Conclusion

What to remember from this talk

- We have presented a convergence theory for algorithm `REGINN` which is based on only 5 features of the underlying inner regularization scheme.
- These features are rather general and are shared by a variety of schemes being so different as Landweber, steepest decent, implicit iteration, and cg-method.

What to remember from this talk

- We have presented a convergence theory for algorithm `REGINN` which is based on only 5 features of the underlying inner regularization scheme.
- These features are rather general and are shared by a variety of schemes being so different as Landweber, steepest decent, implicit iteration, and cg-method.

Thank you for your attention!

GAMM 2010

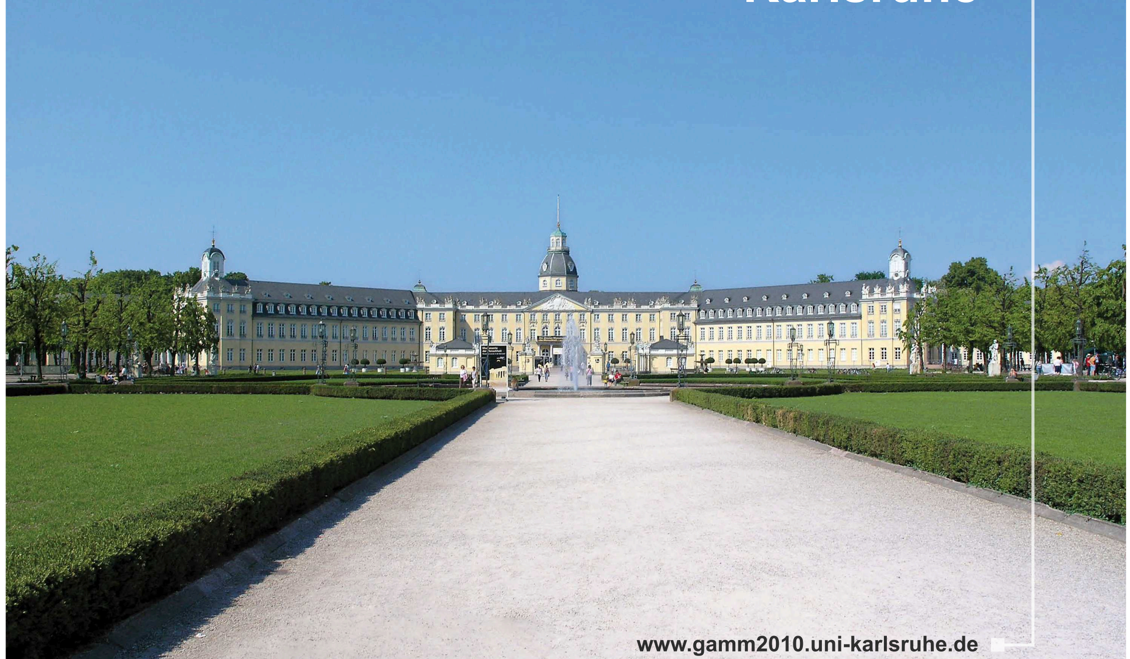
MARCH, 22-26

81st Annual Meeting

of the International Association of Applied Mathematics and Mechanics

at the University of

Karlsruhe



www.gamm2010.uni-karlsruhe.de

Gesellschaft für Angewandte Mathematik und Mechanik

Institut für Wissenschaftliches Rechnen und Mathematische Modellbildung

Universität Karlsruhe (TH)
Forschungsuniversität • gegründet 1825

KIT