Regularizing trust-region and Levenberg-Marquardt approaches for ill-posed nonlinear least squares problems

Stefania Bellavia
Università degli Studi di Firenze
Dipartimento di Ingegneria Industriale

Based on works with Benedetta Morini, Elisa Riccietti

UNIVERSITÀ DEGLI STUDI FIRENZE

OIP2106, Modena, September 19-21 2016.

Outline

- Introduction to iterative regularization methods.
- Zero residual problems:
  - Regularizing Levenberg-Marquardt (LM) methods.
  - Regularizing trust-region (TR) approaches.
  - Numerical tests: LM versus TR
- Small residual problems: Elliptical trust-region methods.
- Open issues and future developments.

Ill-posed problems

Let us consider the following inverse problem: given $F : \mathbb{R}^n \to \mathbb{R}^m$ with $m \geq n$, nonlinear, continuously differentiable and $y \in \mathbb{R}^m$, solve

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \| F(x) - y \|^2.$$

We consider ill-posed problems:

- no finite bounds on the norm of the inverse of $J(x)$ can be used in the analysis;
- the solution does not depend continuously on the data.

Noisy case

In realistic situation only noisy data y^δ are given:

$$\|y - y^\delta\| \leq \delta,$$

where δ is the noise level.

- Applications: Data assimilation, geophysics, seismic inversion, fitting of exponentials, discretization of problems with compact operator
- Classical methods used for well-posed problems are not suitable in this contest.

Need for regularization

Iterative regularization methods

Let $x^\dagger$ be a solution of $\min \frac{1}{2}\|F(x) - y\|^2$.

Iterative regularization methods generate a sequence $\{x_k^\delta\}$. If the process is stopped at iteration $k^*(\delta)$ the method is supposed to guarantee the following properties:

- $x_{k^*(\delta)}^\delta$ is an approximation of $x^\dagger$;
- $\{x_{k^*(\delta)}^\delta\}$ tends to $x^\dagger$ if δ tends to zero;
- local convergence to $x^\dagger$ in the noise-free case.

Regularizing methods for zero residual problems

- Landweber (gradient-type method)[Hanke, Neubauer, Scherzer, 1995,Kaltenbacher, Neubauer, Scherzer, 2008]
- Truncated Newton - Conjugate Gradients [Hanke,1997, Rieder, 2005]
- Iterative Regularizing Gauss-Newton [Bakushinsky, 1992, Blaschke, Neubauer, Scherzer, 1997]
- Levenberg-Marquardt [Hanke,1997,2010,Vogel 1990, Kaltenbacher, Neubauer, Scherzer, 2008]
- Trust region methods [Wang, Yuan 2002,B., Morini, Riccietti 2016]

Most of these methods are analyzed only under local assumptions.

Levenberg-Marquardt method

- Given $x_k^\delta \in \mathbb{R}^n$ and $\lambda_k > 0$, we denote with $J \in \mathbb{R}^{m\times n}$ the Jacobian matrix of F. The step $p_k \in \mathbb{R}^n$ is the minimizer of

$$m_k^{LM}(p) = \frac{1}{2}\|F(x_k^\delta) - y^\delta + J(x_k^\delta)p\|^2 + \frac{1}{2}\lambda_k\|p\|^2;$$

- $p_k = p(\lambda_k)$ is the solution of

$$(B_k + \lambda_k I)p_k = -g_k$$

  with $B_k = J(x_k^\delta)^T J(x_k^\delta)$, $g_k = J(x_k^\delta)^T(F(x_k^\delta) - y^\delta)$;
- The step is then used to compute the new iterate

$$x_{k+1}^\delta = x_k^\delta + p_k.$$

Regularizing LM method for zero residual problems

- The parameter $\lambda_k > 0$ satisfies:

$$\|F(x_k^\delta) - y^\delta + J(x_k^\delta)p(\lambda_k)\| = q\|F(x_k^\delta) - y^\delta\|$$

  with $q \in (0,1)$;

- With noisy data the process is stopped at iteration $k^*(\delta)$ such that $x^\delta_{k^*(\delta)}$ satisfies the discrepancy principle:

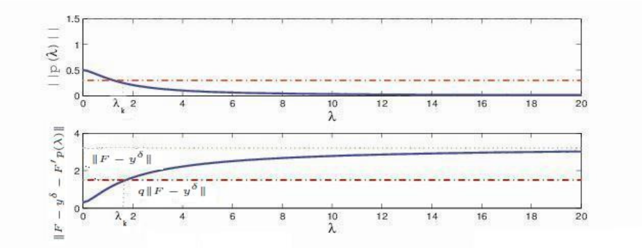
$$\|F(x^\delta_{k^*(\delta)}) - y^\delta\| \leq \tau\delta < \|F(x_k^\delta) - y^\delta\|$$

  for $0 \leq k < k^*(\delta)$ and $\tau > 1$ suitable parameter.

[Hanke, 1997,2010]

The role of the q-condition

- A sufficiently small step is needed in order to prevent to approach the solution of the noisy problem and to leave the region around $x^\dagger$
- The q-condition prevents to take too long steps



$\|p(\lambda)\|$ and $\|F - y^\delta + F'(p(\lambda))\|$ varying λ.

Local analysis

Hypothesis for the local analysis:
Given the starting guess x_0, it exist positive ρ and c such that

- the system $F(x) = y$ is solvable in $B_\rho(x_0)$;
- for $x, \tilde{x} \in B_{2\rho}(x_0)$

$$\|F(x) - F(\tilde{x}) - J(x)(x - \tilde{x})\| \leq c\|x - \tilde{x}\|\|F(x) - F(\tilde{x})\|.$$

Local analysis

Hypothesis for the local analysis:
Given the starting guess x_0, it exist positive ρ and c such that

- the system $F(x) = y$ is solvable in $B_\rho(x_0)$;
- for $x, \tilde{x} \in B_{2\rho}(x_0)$

$$\|F(x) - F(\tilde{x}) - J(x)(x - \tilde{x})\| \leq c\|x - \tilde{x}\|\|F(x) - F(\tilde{x})\|.$$

well posed problems: $\|F(x) - F(\tilde{x}) - J(x)(x - \tilde{x})\| \leq c\|x - \tilde{x}\|^2.$

- Due to the ill-posedness of the problem it is not possible to assume that a finite bound on the inverse of the Jacobian matrix exists.
- The Jacobian may be singular at the solution.

Regularizing properties of the LM method

Choosing λ_k solution of the q-condition

$$\|F(x_k^\delta) - y^\delta + J(x_k^\delta)p(\lambda_k)\| = q\|F(x_k^\delta) - y^\delta\|$$

and stopping the process when the discrepancy principle

$$\|F(x_{k^*(\delta)}^\delta) - y^\delta\| \leq \tau\delta < \|F(x_k^\delta) - y^\delta\|$$

is satisfied, Hanke proves that:

- With exact data ($\delta = 0$): local convergence to $x^\dagger$,
- With noisy data ($\delta > 0$): Choosing x_0 close to $x^\dagger$ the discrepancy principle is satisfied after a finite number of iterations $k^*(\delta)$ and $\{x_{k^*(\delta)}^\delta\}$ converges to a solution of $F(x) = y$ if δ tends to zero.

Regularizing method.