Levenberg-Marquardt methods for the solution of noisy nonlinear least squares problems

Elisa Riccietti

Università degli Studi di Firenze
Dipartimento di Matematica e Informatica 'Ulisse Dini'
National Polytechnic Institute (INP), Toulouse

Relatore: Stefania Bellavia, Tutor estero: Serge Gratton

Seminario di passaggio di anno, 27/09/2017.

Key ingredients

Nonlinear least squares problems

Given $R : \mathbb{R}^n \to \mathbb{R}^m$, $m \geq n$, nonlinear, continuously differentiable solve

$$\min_{x \in \mathbb{R}^n} \Phi(x) = \frac{1}{2}\|R(x)\|^2.$$

Levenberg-Marquardt method

It is an iterative method that builds the sequence of solution approximations as $x_{k+1} = x_k + p_k$ where p_k is the solution of:

$$\min_{p \in \mathbb{R}^n} m_k^{LM}(p) = \frac{1}{2}\|R(x_k) + J(x_k)p\|^2 + \frac{1}{2}\lambda_k\|p\|^2$$

where J is the Jacobian matrix of R and $\lambda_k \geq 0$ is a regularization parameter.

Classical Levenberg-Marquardt method

- Given $x_k \in \mathbb{R}^n$ and $\lambda_k \geq 0$, the step $p_k \in \mathbb{R}^n$ is the minimizer of

$$m_k^{LM}(p) = \frac{1}{2}\|R(x_k) + J(x_k)p\|^2 + \frac{1}{2}\lambda_k\|p\|^2.$$

- p_k is the solution of

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

  with $B_k = J(x_k)^T J(x_k)$, $g_k = J(x_k)^T R(x_k)$.
- Set $\Phi(x) = \frac{1}{2}\|R(x)\|^2$, and compute

$$\rho_k(p_k) = \frac{\Phi(x_k) - \Phi(x_k + p_k)}{m_k^{LM}(0) - m_k^{LM}(p_k)}.$$

- Given $\eta \in (0, 1)$:
  - If $\rho_k < \eta$ then set $\lambda_{k+1} > \lambda_k$ and $x_{k+1} = x_k$.
  - If $\rho_k \geq \eta$ then set $\lambda_{k+1} \leq \lambda_k$ and $x_{k+1} = x_k + p_k$.

Outline

$$\min_{x\in\mathbb{R}^n} \Phi(x) = \frac{1}{2}\|R(x)\|^2.$$

The thesis is divided into two parts:

- I part: Ill-posed problems with $R(x) = F(x) - y$ for given data y. We assume to have at disposal just noisy data y^δ. The noise is fixed and arises from measurements errors. AIM: design stable methods for their solution.
- II part: Large scale problems with noisy function and gradient, $R(x) = F_\delta(x)$ noisy approximation to $F(x)$. The approximation can be improved reducing the noise level. AIM: design fast methods for the solution of the unperturbed problem for noise level converging to zero.

I part: Ill-posed least squares problems

I part: Ill-posed least squares problems

Let us consider the following least squares problem: given $\mathcal{X}, \mathcal{Y}$ Hilbert spaces, $F : \mathcal{X} \to \mathcal{Y}$, nonlinear, continuously differentiable and $y \in \mathcal{Y}$, solve

$$\min_{x \in \mathcal{X}} \|F(x) - y\|^2.$$

Definition

The problem is well-posed if:

1. $\forall y \in \mathcal{Y}$ it exists a solution $x \in \mathcal{X}$,
2. the solution is unique,
3. property of stability holds.

The problem is ill-posed if one or more of the previous properties do not hold.

Ill-posed problems

- Let us consider problems of the form

$$\min_{x \in \mathbb{R}^n} \|F(x) - y\|^2, \qquad x \in (\mathbb{R}^n, \|\cdot\|_2),\ y \in (\mathbb{R}^m, \|\cdot\|_2),$$

  with $F : \mathbb{R}^n \to \mathbb{R}^m$ and $m \geq n$, arising from the discretization of an ill-posed problem.
- In a realistic situation the data y are affected by noise, we have at disposal only y^δ such that:

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

  for some positive δ .
- We can handle only a noisy problem:

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

Need for regularization

- As stability does not hold, the solutions of the original problem do not depend continuously on the data.
  $\Longrightarrow$ The solutions of the noisy problem may not be meaningful approximations of the original problem solutions.

Need for regularization

- As stability does not hold, the solutions of the original problem do not depend continuously on the data.
  $\Longrightarrow$ The solutions of the noisy problem may not be meaningful approximations of the original problem solutions.
- For ill-posed problems there are no finite bounds on the inverse of the Jacobian of F around a solution of the original problem.

Need for regularization

- As stability does not hold, the solutions of the original problem do not depend continuously on the data.
  $\Longrightarrow$ The solutions of the noisy problem may not be meaningful approximations of the original problem solutions.
- For ill-posed problems there are no finite bounds on the inverse of the Jacobian of F around a solution of the original problem.
- Classical methods used for well-posed systems are not suitable in this contest.

$$\Downarrow$$

Need for regularization.

Outline

1) Zero-residual problems: $F(x) = y^\delta$

It exists $x^\dagger$ such that $F(x^\dagger) = y$. We propose a regularizing trust-region approach, able to find an approximation to a solution of the unperturbed problem.

2) Non-zero residual problems: $\min_{x \in \mathbb{R}^n} \|F(x) - y^\delta\|^2$

It does not exist $x^\dagger$ such that $F(x^\dagger) - y = 0$.
We extend the trust-region approach designed for zero-residual problem to small residual problems.

Trust-region methods

Trust region methods falls into the class of Levenberg-Marquardt methods.

Levenberg-Marquardt - Trust region

- LM: $\min_p m_k^{LM}(p) = \frac{1}{2}\|F(x_k) - y + J(x_k)p\|^2 + \frac{\lambda_k}{2}\|p\|^2$
- TR: $\min_p m_k^{TR}(p) = \frac{1}{2}\|F(x_k) - y + J(x_k)p\|^2,$
  s.t. $\|p\| \leq \Delta_k$

It is possible to prove that for TR p_k solves

$$(B_k + \lambda_k I)p_k = -g_k, \quad B_k = J(x_k)^T J(x_k),\ g_k = J(x_k)^T (F(x_k) - y)$$

for some $\lambda_k \geq 0$ such that

$$\lambda_k(\|p_k\| - \Delta_k) = 0.$$

$\Rightarrow$ Trust-region methods are Levenberg-Marquardt methods!

Trust-region methods

- At each iteration the step is accepted if it provides sufficient decrease in the objective function $\Phi(x) = \frac{1}{2}\|F(x) - y^\delta\|^2$ and the trust region radius is updated.
- The update is based on the ratio between actual and predicted reduction:

$$\rho_k(p_k) = \frac{\Phi(x_k) - \Phi(x_k + p_k)}{m_k^{TR}(0) - m_k^{TR}(p_k)}.$$

- Given $\eta \in (0, 1)$:
  - If $\rho_k < \eta$ then set $\Delta_{k+1} < \Delta_k$ and $x_{k+1} = x_k$.
  - If $\rho_k \geq \eta$ then set $\Delta_{k+1} \geq \Delta_k$ and $x_{k+1} = x_k + p_k$.

Zero-residual problems

We consider

$$F(x) = y^{\delta},$$

with δ fixed noise level, and let $x^{\dagger}$ be a solution of $F(x) = y$.

Iterative regularization methods

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:

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

Standard trust-region

Standard trust-region

The step p_k solves

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

for some $\lambda_k \geq 0$ such that

$$\lambda_k(\|p_k\| - \Delta_k) = 0.$$

- B_k is ill-conditioned.
- In trust-region methods the trust region is eventually inactive: $\|p_k\| < \Delta_k \rightarrow \lambda_k = 0$.
- It is not a regularization method!

How to obtain a regularizing method?

Noisy problem

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

Exact problem

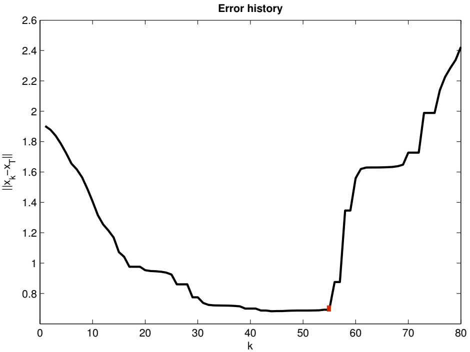
$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|F(x) - y\|^2$$

1. stopping criterion
2. small steps

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

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

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

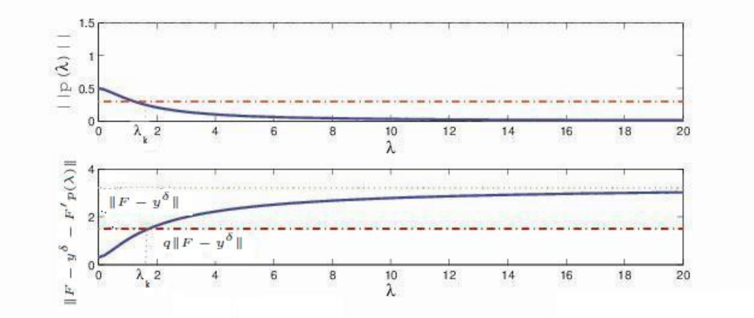


SEMI CONVERGENCE

Plot of the error $\|x^\delta_k - x^\dagger\|$ versus iteration number.

Regularizing trust-region

2) q-condition: $\|F(x_k^\delta) - y^\delta + J(x_k^\delta)p\| \geq q\|F(x_k^\delta) - y^\delta\|$, $q \in (0,1)$



$\rightarrow$ If $\Delta_k \leq \frac{1-q}{\|B_k\|}\|g_k^\delta\|$ then p_k satisfies the q-condition and the trust region is active.

- **Assumption 1** For index $\bar{k}$ it exist positive ρ and c such that
  1. the system $F(x) = y$ is solvable in $B_\rho(x^\delta_{\bar{k}})$;
  2. for $x, \tilde{x} \in B_{2\rho}(x^\delta_{\bar{k}})$

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

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

- **Assumption 2**: It exists positive K_J such that

$$\|J(x)\| \leq K_J$$

for all $x \in \mathcal{L} = \{x \in \mathbb{R}^n \ \ s.t. \ \ \Phi(x) \leq \Phi(x_0)\}$.

Theoretical results

Lemma

The method generates a sequence $\{x_k^\delta\}$ such that:
1) the trust-region is active, i.e. $\lambda_k > 0$,
2) error decreases monotonically: $\|x_{k+1}^\delta - x^\dagger\| < \|x_k^\delta - x^\dagger\|$,
for $k \geq \bar{k}$, with $\bar{k} < k^*(\delta)$ for noisy data.

Theorem

If $\delta = 0$ the sequence $\{x_k\}$ converges to a solution x^* of $F(x) = y$ such that $\|x^* - x^\dagger\| \leq \rho$.
If $\delta > 0$ the discrepancy principle is satisfied after a finite number of iterations $k^*(\delta)$ and the sequence $\{x_{k^*(\delta)}^\delta\}$ converges to a solution of $F(x) = y$ if δ tends to zero.

$\rightarrow$ **Regularizing method**, [S. Bellavia, B. Morini, E. R., COAP, 2016].

Test problems

- Four nonlinear ill-posed systems arising from the discretization of nonlinear first-kind Fredholm integral equation are considered, they model gravimetric and geophysics problems:

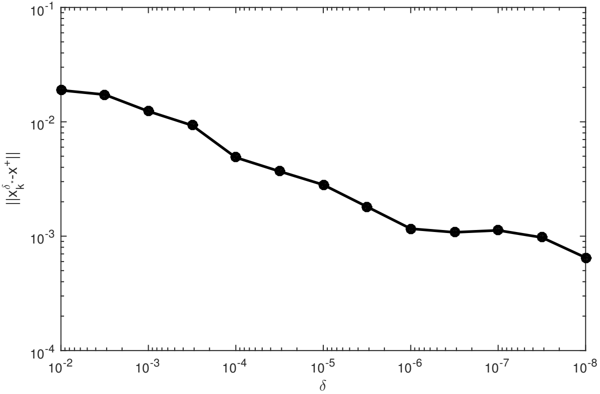
$$\int_0^1 k(t,s,x(s))ds = y(t), \qquad t \in [0,1],$$

P1,**P2**, [Vogel, 1990], **P3**,**P4** [Kaltenbacher,2007];

- Their kernel is of the form

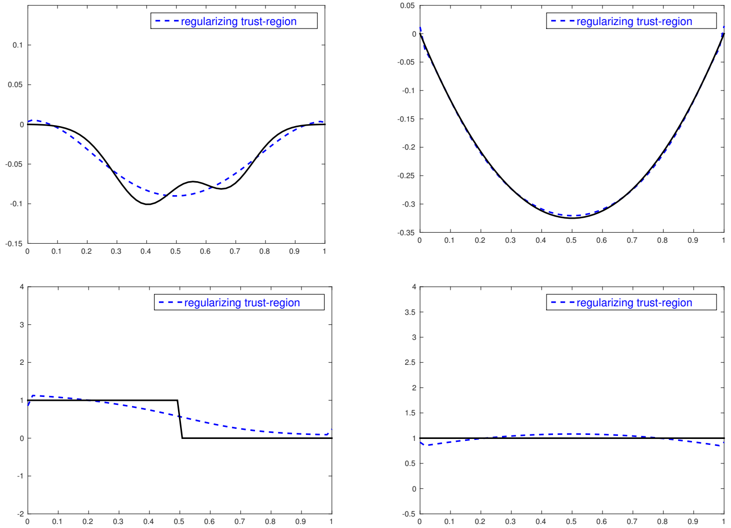
$$\begin{aligned} k(t,s,x(s)) &= log\left(\frac{(t-s)^2+H^2}{(t-s)^2+(H-x(s))^2}\right); \\ k(t,s,x(s)) &= \frac{1}{\sqrt{1+(t-s)^2+x(s)^2}}; \end{aligned}$$

Regularizing properties of the method.



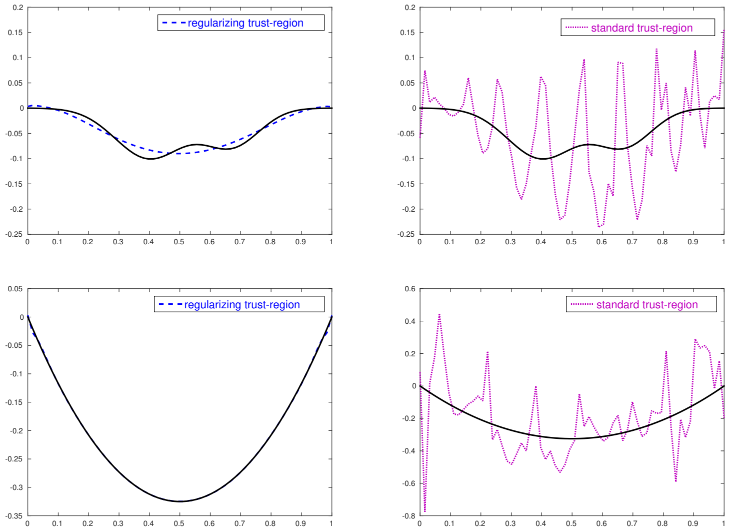
Semilogarithmic plot of the error $||x^{\delta}_{k^*(\delta)} - x^{\dagger}||$ as a function of the noise level δ.

Computed solution approximations



Blue: regularizing TR. **Solid line:** solution of the original problem

Comparison between regularizing and standard trust-region



Non-zero residual problems

We consider non-zero residual problems:
it does not exist x such that $F(x) - y = 0$, but it exists $x^{\dagger}$ local minimum of the problem.

Small residual problems

- We extend the approach for zero-residual problems to small residual problems. → We propose an **elliptical trust-region approach.**
- Let us assume that J is full rank, but ill-conditioned, with ill-conditioning due to smallest singular value close to zero, as often happens in these applications.

At a generic iteration k, given $\Delta_k > 0$, the following problem is solved:

$$\min_p m_k(p) := \frac{1}{2}\|F(x_k^\delta) - y^\delta + J(x_k^\delta)p\|^2,$$
$$\text{s.t. } \|(B_k)^{-\frac{1}{2}} p\| \leq \Delta_k.$$

Regularizing properties

To maintain the regularizing properties of the trust-region approach we assume equivalent conditions on the gradient instead on the function.

1. **discrepancy principle** :

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

2. **q-condition**:

$$\|J(x^\delta_k)^T(F(x^\delta_k) - y^\delta + J(x^\delta_k)p_k)\| \geq q\|J(x^\delta_k)^T(F(x^\delta_k) - y^\delta)\|$$

If $\Delta_k \leq \frac{1-q}{\|B_k\|^2}\|(B_k)^{1/2}g^\delta_k\|$ then p_k satisfies the q-condition and the trust-region is active.

Convergence analysis

- **Assumption1**: there exists $\bar{k}$ s.t. a solution exists in $B_\rho(x_{\bar{k}})$ and for $x, \tilde{x} \in B_{2\rho}(x_{\bar{k}})$

$$\|\nabla f(\tilde{x}) - \nabla f(x) - J(x)^T J(x)(\tilde{x} - x)\| \leq (c\|\tilde{x} - x\| + \sigma)\|\nabla f(x) - \nabla f(\tilde{x})\|.$$

$$\nabla^2 f(x) = J(x)^T J(x) + S(x) = J(x)^T J(x) + \sum_{j=1}^{m} (F_j(x) - y_j)\nabla^2 F_j(x).$$

- **Assumption2**: $\|S(x^\dagger)\| \leq \sigma < q < 1$ (small residual problems)

$$\Downarrow$$

Regularizing method.

Convergence analysis

- Let $\delta = 0$. Under Assumptions 1,2 the sequence $\{x_k\}$ generated converges to a stationary point x^* such that $\|x^* - x^\dagger\| \leq \rho$.
- Let $\delta > 0$. Under Assumptions 1,2 the iterates satisfy the discrepancy principle after a finite number $k_*(\delta)$ of iterations. Moreover the sequence $\{x^\delta_{k_*(\delta)}\}$ converges to a stationary point whenever δ tends to zero.

Numerical results

1. **P1**: We want to reconstruct c in the 2D-elliptic problem

$$-\Delta u + cu = \hat{f} \text{ in } \Omega = (0,1) \times (0,1)$$
$$u = \hat{g} \text{ on } \partial\Omega$$

   from the knowledge of u in Ω, $\hat{f} \in L^2(\Omega)$, $\hat{g}$ the trace of a function in $H^2(\Omega)$. If $F : D(F) \to L^2(\Omega)$ is the operator mapping parameter c to the solution u we solve

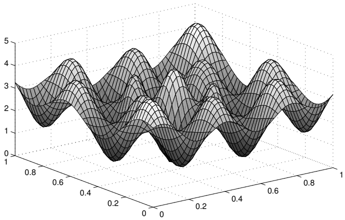
$$\min_c \frac{1}{2}\|F(c) - \tilde{u}\|^2$$

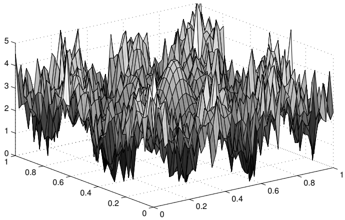
   $\tilde{u}$ measured values of u.

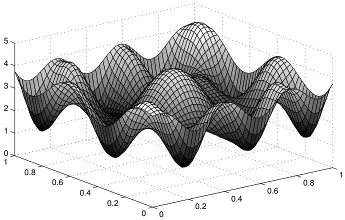
2. **P2**: Reconstruct the conductivity x of the soil from measurements $b = (b_1, \dots, b_m)^T$ at different heights $h_i, i = 1, \dots, m$:

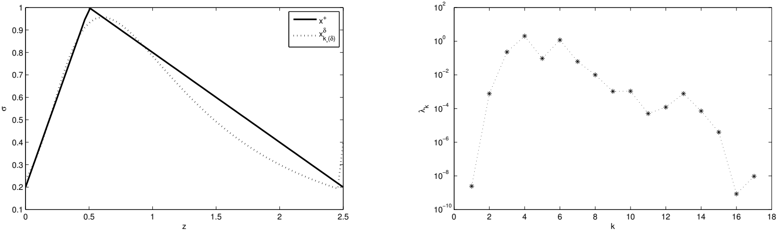
$$\min_x \frac{1}{2}\|m(x) - b\|^2.$$

Numerical tests on problem P1, $\delta = 1.e-2$









(a) (b)

Figure: (a) plot of the true solution $x^\dagger$ and of the computed solution $x^\delta_{k^*(\delta)}$ for $\delta = 10^{-2}$, (b) regularization parameters λ_k.

II part: Large scale noisy problems

II part: Large scale noisy problems

- 6-months collaboration with S. Gratton, INP-ENSEEIHT, Toulouse.
- We consider problems for which the objective function is expensive to evaluate.
- We want to recover the solution of the problem relying on cheap approximations to the objective function:

$$\min_x \frac{1}{2}\|F_\delta(x)\|^2 \qquad\qquad \|F_\delta(x) - F(x)\| \leq \delta.$$

- It is possible to improve the approximation quality decreasing the noise level δ during the optimization process.
- Generally we will deal with large-scale non-zero residual problems.

Levenberg-Marquardt method

- At each iteration we consider

$$\min_x \Phi_{\delta_k}(x) = \frac{1}{2}\|F_{\delta_k}(x)\|^2 \qquad \|F_{\delta_k}(x) - F(x)\| \leq \delta_k.$$

- At each iteration we have to solve a noisy linear systems of the form:

$$(J_{\delta_k}(x_k^{\delta_k})^T J_{\delta_k}(x_k^{\delta_k}) + \lambda_k I)p_k = -g_{\delta_k}(x_k^{\delta}).$$

- We have to compute:

$$\rho_k^{\delta_k}(p_k) = \frac{\Phi_{\delta_k}(x_k^{\delta_k}) - \Phi_{\delta_k}(x_k^{\delta_k} + p_k)}{m_k(0) - m_k(p_k)}.$$

- If the noise is too high the reduction in Φ_{δ_k} can be just an effect of the presence of the noise.

Noise control

Let

$$\delta_k \leq \frac{1}{2}\lambda_k^{\alpha}\|p\|^2,$$

for $\alpha \in (\frac{1}{2}, 1]$. If

$$\rho_k^{\delta_k}(p_k) = \frac{\Phi_{\delta_k}(x_k^{\delta_k}) - \Phi_{\delta_k}(x_k^{\delta_k} + p_k)}{m_k(0) - m_k(p_k)} > \eta$$

then also

$$\rho_k(p_k) = \frac{\Phi(x_k^{\delta_k}) - \Phi(x_k^{\delta_k} + p_k)}{m_k(0) - m_k(p_k)} > \eta.$$

$\rightarrow$ True reduction in the noise-free objective function Φ

Algorithm : k-th iteration of regularizing Levenberg-Marquardt

Given $\alpha \in \left(\frac{1}{2}, 1\right]$, δ_0, $\eta_1 \in (0,1)$, $\eta_2 > 0$, $\lambda_{\max} > \lambda_{\min} > 0$, $\gamma > 1$, x_0 and $\lambda_{\max} > \lambda_0 \geq \lambda_{\min}$.

Compute $f^{\delta_0}(x_0)$. For $k = 0, 1, 2, \ldots$

1. Compute a solution p_k of the LM subproblem.
2. If $\delta_k \leq \frac{1}{2}\lambda_k^\alpha \|p_k\|^2$, compute $\Phi_{\delta_k}(x_k^{\delta_k} + p_k)$, else reduce δ_k and go back to 1.
3. Compute

$$\rho_k^{\delta_k}(p_k) = \frac{\Phi_{\delta_k}(x_k^{\delta_k}) - \Phi_{\delta_k}(x_k^{\delta_k} + p_k)}{m_k(0) - m_k(p_k)}.$$

3.1 If $\rho_k^{\delta_k}(p_k) \geq \eta_1$, then set $x_{k+1}^{\delta_k} = x_k^{\delta_k} + p_k$ and

$$\lambda_{k+1} = \begin{cases} \min\{\gamma\lambda_k, \lambda_{\max}\} & \text{if } \|g_{\delta_k}(x_k^\delta)\| < \eta_2/\lambda_k, \\ \max\{\lambda_k, \lambda_{\min}\} & \text{if } \|g_{\delta_k}(x_k^\delta)\| \geq \eta_2/\lambda_k. \end{cases}$$

3.2 Otherwise set $x_{k+1}^{\delta_k} = x_k^{\delta_k}$, $\lambda_{k+1} = \gamma\lambda_k$.

Inexact step

Large-scale problems: approximate solution of LM subproblem

p provides the sufficient Cauchy decrease:

$$m_k(0) - m_k(p_k) \geq \frac{\theta}{2} \frac{\|g_{\delta_k}(x_k^\delta)\|^2}{\|J_{\delta_k}(x_k^\delta)\|^2 + \lambda_k}, \qquad \theta > 0.$$

The Levenberg-Marquardt step computed as

$$(J_{\delta_k}(x_k^{\delta_k})^T J_{\delta_k}(x_k^{\delta_k}) + \lambda_k I)p_k = -g_{\delta_k}(x_k^\delta) + r_k$$

for a residual r_k satisfying $\|r_k\| \leq \epsilon_k \|g_{\delta_k}(x_k^\delta)\|$, with ϵ_k small enough achieves the Cauchy decrease.

Assumptions

Let $\mathcal{L} = \{x \mid f(x) \leq f(x_0)\}$, $x^\dagger$ a local minimum of function f.

- Assumption 1: It exists $K > 0$ such that $\|J(x)\| \leq K$ for all $x \in \mathcal{L}$.
- Assumption 2: f has Lipschitz continuous gradient on $\mathcal{L}$:
  $\|g(x) - g(y)\| \leq L\|x - y\|$ for all $x, y \in \mathcal{L}$.
- Assumption 3: Let H the Hessian matrix of function f, $H(x^\dagger) \succeq 0$, H is Lipschitz continuous: $\|H(x) - H(y)\| \leq M\|x - y\|$ for all $x, y \in \mathcal{L}$, and let $0 < l \leq L < \infty$ such that $l\, I_n \preceq H(x^\dagger) \preceq L\, I_n$ with I_n the identity matrix of size n.

Theoretical results

- Global convergence: Let Assumptions 1 and 2 hold. Then

$$\lim_{k\to\infty} \delta_k = 0, \qquad \lim_{k\to\infty} \|g_{\delta_k}(x_k^\delta)\| = 0.$$

- Parameters $\lambda_{\mathbf{k}}$: It exists $\bar{k} > 0$ such that $\lambda_k = \lambda_{\max}$ for all $k \geq \bar{k}$.
- Asymptotic step behaviour Let p_k satisfy the Cauchy decrease. Then

$$\lim_{k\to\infty} (p_k^{LM})_i + \frac{\theta}{K^2 + \lambda_k}(g_{\delta_k}(x_k^\delta))_i = 0 \quad \text{for} \quad i = 1, \dots, n,$$

  If λ_k is large enough p_k tends to a steepest descent step with step-length $\frac{1}{\lambda_{\max}}$.
- Local Convergence perturbed steepest descent: Let Assumptions 1,2 and 3 hold and let $p_k^{SD} = -\frac{1}{\lambda_{\max}} g_{\delta_k}(x_k^\delta)$ with $\frac{1}{\lambda_{\max}} < \frac{1}{L}$. If the starting guess x_0 is close enough to $x^\dagger$, the method converges locally.

Numerical results

Nonlinear wave equation:

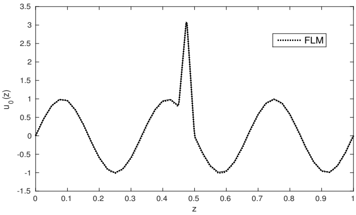
$$\frac{\partial^2 u(z,t)}{\partial t^2} - \frac{\partial^2 u(z,t)}{\partial z^2} + \mu e^{\nu u} = 0,$$
$$u(0,t) = u(1,t) = 0, u(z,0) = u_0(z),$$
$$\frac{\partial u(z,0)}{\partial t} = 0,\ 0 \le t \le T,\ 0 \le z \le 1.$$

We look for the initial state $u_0(z)$, from the knowledge of observations $u(z_i, t_j)$, $t_j > 0$. Data assimilation problem:

$$\min_{x \in \mathbb{R}^n} \frac{1}{2}\|x - x_b\|^2_{B^{-1}} + \frac{1}{2}\sum_{j=0}^{N_t} \|H_j(x(t_j)) - y_j\|^2_{R_j^{-1}}$$

- $\|x\|^2_M = x^T M x$ for a symmetric positive definite matrix M,
- $x_b \in \mathbb{R}^n$ is the background vector (a priori estimate)
- $y_j \in \mathbb{R}^{m_j}$ is the vector of observations at time t_j, $m_j \le n$.
- H_j is the operator modelling the observation process at t_j
- $x(t_j)$ the state vector, solution of the nonlinear model at time t_j.

Numerical results



	All samples	Subsampled
it	9	12
$\mathbf{cost_f}$	10	3
$\mathbf{cost_p}$	67	15
RMSE	1.2e-2	3.8e-2
$\mathbf{save_f}$		67%
$\mathbf{save_p}$		78%

Research Outputs

Articles related to the thesis:

- S.Bellavia, B.Morini, E.Riccietti, *On an adaptive regularization for ill-posed nonlinear systems and its trust-region implementation* (COAP, 2016).
- S.Bellavia, E.Riccietti, *On non-stationary Tikhonov procedures for ill-posed nonlinear least squares problems*, submitted to Inverse Problems.
- S.Bellavia, S.Gratton, E.Riccietti, *A Levenberg-Marquardt method for large nonlinear least squares problems with noisy functions and gradients*, submitted to Numerische Mathematik.

Other articles:

- E.Riccietti, J.Bellucci, M.Checcucci, M.Marconcini, A.Arnone, *Support Vector Machine classification applied to the parametric design of centrifugal pumps*, (ENGOPT, 2017).
- E.Riccietti, S.Bellavia, S.Sello, *Numerical methods for optimization problems arising in energetic districts*, (ECMI proceeding, 2016).
- E.Riccietti, S.Bellavia, S.Sello, *Sequential Linear Programming and Particle Swarm Optimization for the optimization of energy districts*, submitted to Engineering Optimization.