Inexact Newton Combined with Gradient Methods in Banach Spaces

Fábio Margotti

Universidade Federal de Santa Catarina

Rio de Janeiro, August 31th 2016

Outline

Ill-Posed Problems

Inverse Problems

Find the CAUSE of a phenomenon from partial knowledge of the EFFECT produced by it.

Mathematically: Find x satisfying

$$F(x) = y,$$

$F : X \supset D(F) \to Y$ operates between *Banach* spaces.

Definition (Well-posedness (Hadamard))

1. Existence
2. Uniqueness
3. Stability

Regularization Methods

Noisy data: Find x satisfying

$$F(x) = y,$$

having

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

(Regularization Property)

For each pair (y^δ, δ) *find a vector*

$$x_\delta \approx x^+$$

such that

$$x_\delta \to x^+ \qquad as \qquad \delta \to 0.$$

Difficulties in Banach Spaces

$$As = b, \qquad A : X \longrightarrow Y \text{ linear and bounded}$$

Hilbert spaces:

$\varphi(s) := \frac{1}{2}\|As - b\|^2$ is differentiable,

$\psi_k := \nabla\varphi(s_k) = A^*(As_k - b) \in X$,

$s_{k+1} := s_k - \lambda\psi_k$, with $\lambda > 0$.

Banach spaces:

$\varphi(s) := \frac{1}{r}\|As - b\|^r$, $r > 1$, is subdifferentiable,

$\psi_k \in \partial\varphi(s_k) \subset X^*$,

$$j : X \longrightarrow X^* \quad \text{and} \quad j^* : X^* \longrightarrow X \quad \text{are necessary.}$$

Duality Mapping

$$f(x) := \frac{1}{p}\|x\|^p, \qquad p > 1$$

Convexity: related to convexity of f

Smoothness: related to differentiability of f

Definition (Duality Mapping)

$$J_p(x) := \partial f(x)$$

Selection: $j_p : X \longrightarrow X^*, \quad j_p(x) \in J_p(x)$

$$\langle j_p(x), y\rangle \leq \|x\|^{p-1}\|y\| \qquad \text{and} \qquad \langle j_p(x), x\rangle = \|x\|^p$$

Bregman Distance

Polarization Identity:

$$\frac{1}{2}\|x-y\|^2 = \frac{1}{2}\|x\|^2 - \langle y, x\rangle + \frac{1}{2}\|y\|^2$$

Definition (Bregman Distance)

$$\Delta_p(x,y) := \frac{1}{p}\|x\|^p - \langle j_p(y), x\rangle + \frac{1}{p^*}\|j_p(y)\|^{p^*}$$

Δ_p is not a metric

Definition

$$X \text{ is } p\text{-convex} \Leftrightarrow \Delta_p(x,y) \gtrsim \frac{1}{p}\|x-y\|^p \text{ for all } x,y \in X$$

$$X \text{ is } p\text{-smooth} \Leftrightarrow \Delta_p(x,y) \lesssim \frac{1}{p}\|x-y\|^p \text{ for all } x,y \in X$$

Smoothness/Convexity of Power Type

$$X \ p\text{−smooth} \Rightarrow X \text{ uniformly smooth} \Rightarrow X \text{ smooth}$$

$$X \ p\text{−convex} \Rightarrow X \text{ uniformly convex} \Rightarrow X \text{ strictly convex}$$

Examples

1. The spaces $L^p(\Omega)$, $W^{n,p}(\Omega)$ and $\ell^p(\mathbb{R})$, $1 < p < \infty$, are

$$\max\{p, 2\}-\text{convex} \quad \text{and} \quad \min\{p, 2\}-\text{smooth}$$

2. Hilbert spaces are 2−convex and 2−smooth

3. $J_p : L^p(\Omega) \longrightarrow L^{p^*}(\Omega), \quad 1 < p < \infty$

$$J_p(g) = |g|^{p-1}\text{sgn}(g)$$

Assumptions

1. $B := B_\rho(x^+, \Delta_p) \subset D(F) \quad \text{with} \quad F(x^+) = y$

2. F F-differentiable and $\|F'(x)\| \leq M, \ \forall x \in B$

3 (TCC). $\exists\ 0 \leq \eta < 1$ such that $\forall x, w \in B$

$$\|F(x) - F(w) - F'(w)(x - w)\| \leq \eta \|F(x) - F(w)\|$$

4. X is loc. unif. smooth and $p-$convex

REGularization INexact Newton

Algorithm (REGINN (A. Rieder, 1999))

Input: $x_0, F, y^\delta, \delta, \mu, \tau$

$n = 0,$

Repeat n

$s_{n,0} = 0, \quad A_n = F'(x_n), \quad b_n^\delta = y^\delta - F(x_n), \quad k = 0,$

Repeat k

Choose $\lambda_{n,k} > 0$ *and* $\psi_{n,k} \in A_n^* J_r(A_n s_{n,k} - b_n^\delta)$

$J_p(x_n + s_{n,k+1}) = J_p(x_n + s_{n,k}) - \lambda_{n,k}\psi_{n,k}$

Until $\left\| A_n s_{n,k} - b_n^\delta \right\| < \mu \left\| b_n^\delta \right\|$

$x_{n+1} = x_n + s_{n,k},$

Until $\left\| b_n^\delta \right\| \leq \tau\delta$

Decreasing Error

$$J_p(z_{k+1}) := J_p(z_k) - \lambda\psi_k, \quad \psi_k \in \partial\varphi(s_k) = A^* J_r(As_k - b)$$

$e := x^+ - x_n$

From Three Points Identity,

$$\Delta_p\left(x^+, z_{k+1}\right) - \Delta_p\left(x^+, z_k\right) = \Delta_p(z_k, z_{k+1}) + \left\langle J_p\left(z_{k+1}\right) - J_p\left(z_k\right), z_k - x^+\right\rangle$$

But,

$$\begin{aligned}\Delta_p\left(z_k, z_{k+1}\right) &= \Delta_{p^*}\left(J_p\left(z_{k+1}\right), J_p\left(z_k\right)\right)\\ &\le C_{p^*}\left\|J_p\left(z_{k+1}\right) - J_p\left(z_k\right)\right\|^{p^*} = C_{p^*}\lambda^{p^*}\left\|\psi_k\right\|^{p^*}\end{aligned}$$

and

$$\begin{aligned}\left\langle J_p\left(z_{k+1}\right) - J_p\left(z_k\right), z_k - x^+\right\rangle &= -\lambda\left\langle A^* j_r\left(As_k - b\right), s_k - e\right\rangle\\ &\le -\lambda\left\{\|As_k - b\|^r - \|As_k - b\|^{r-1}\|Ae - b\|\right\}\\ &\le -\lambda\left(1 - \frac{\eta}{\mu}\right)\|As_k - b\|^r\end{aligned}$$

Decreasing Error

Thus

$$\Delta_p\left(x^+, z_{k+1}\right) - \Delta_p\left(x^+, z_k\right) \leq C_{p^*}\lambda^{p^*}\left\|\psi_k\right\|^{p^*} - \lambda\left(1 - \frac{\eta}{\mu}\right)\left\|As_k - b\right\|^r$$

If

$$\lambda < \lambda_{\max} := C_0 \frac{\|As_k - b\|^{r(p-1)}}{\|\psi_k\|^p}, \qquad C_0 := \left(\frac{1-\eta/\mu}{C_{p^*}}\right)^{p-1} > 0,$$

Then

$$\Delta_p\left(x^+, z_{k+1}\right) < \Delta_p\left(x^+, z_k\right),$$

Which implies

$$\Delta_p\left(x^+, x_{n+1}\right) = \Delta_p\left(x^+, z_{k_n}\right) < \Delta_p\left(x^+, z_0\right) = \Delta_p\left(x^+, x_n\right),$$

Results

1. If λ_{max} is a bit smaller then

a) $0 < \lambda \leq \lambda_{\text{max}} \Rightarrow \Delta_p(x^+, z_{k+1}) - \Delta_p(x^+, z_k) \lesssim -\lambda \|As_k - b\|^r$

b) The same holds if $\delta > 0$ and τ large enough

2. $\lambda \in [\lambda_{\text{min}}, \lambda_{\text{max}}], \quad \lambda_{\text{min}} := c\|As_k - b^\delta\|^t$ implies

a) $k_n < \infty$

b) $N(\delta) < \infty$

c) $(x_n)_{0 \leq n \leq N(\delta)}$ uniformly bounded

Results

3. η small enough $\Rightarrow \|y^\delta - F(x_{n+1})\| \leq \Lambda \|y^\delta - F(x_n)\|, \quad \Lambda < 1$

and

$$N(\delta) \leq \log_\Lambda \left(\frac{\tau}{\|y^\delta - F(x_0)\|} \delta \right) + 1$$

4. $\delta = 0$ implies

$$(x_n) \text{ converges to a solution as } n \longrightarrow \infty$$

5. Y loc. unif. smooth and $\left(x_n^{\delta_j} \right)_{0 \leq n \leq N(\delta_j)}$ fixed with $\delta_j \longrightarrow 0$
$\Rightarrow \exists$ a noiseless sequence (x_n) s.t.

$$\forall M \in \mathbb{N}, \quad \exists \text{ a subsequence } \left(\delta_{j_m} \right)_m \text{ satisfying}$$

$$x_n^{\delta_{j_m}} \longrightarrow x_n \quad \text{as} \quad m \longrightarrow \infty, \quad 0 \leq n \leq M$$

Results

6. Regularization Property: If $\delta_j \longrightarrow 0$, then

a) Each subsequence of $\left(x_{N(\delta_j)}^{\delta_j}\right)_j$ has itself a subsequence which converges to a solution

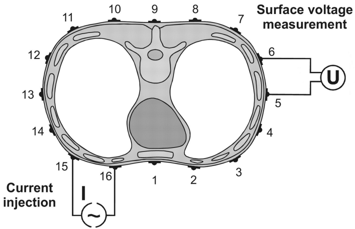
b) If the solution is unique,

$$x_{N(\delta_j)}^{\delta_j} \longrightarrow x^+ \quad \text{as} \quad j \longrightarrow \infty$$

c) If $\left(\lambda_{n,k}^{\delta_j}\right)_j$ converges, then

$$\left(x_{N(\delta_j)}^{\delta_j}\right)_j \text{ converges to a solution as } j \longrightarrow \infty$$

Electrical Impedance Tomography



[I. Frerichs et al. 2001]

Neumann-to-Dirichlet operator

$$\Lambda_\sigma : \mathbb{R}^d \longrightarrow \mathbb{R}^d, \qquad I \mapsto U$$

Inverse Problem: Find $\sigma \in L^\infty(\Omega)$ satisfying

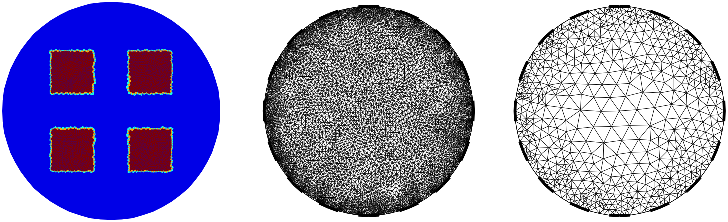
$$F(\sigma) := \Lambda_\sigma,$$

having

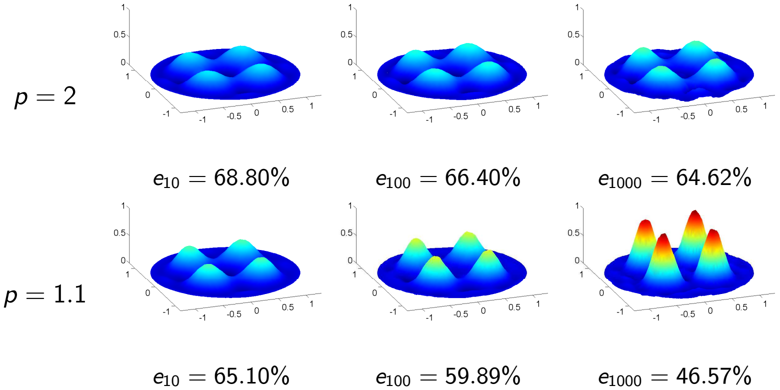
$$\|\Lambda_\sigma^\delta - \Lambda_\sigma\|_{\mathcal{L}(\mathbb{R}^d,\mathbb{R}^d)} \leq \delta.$$

Electrical Impedance Tomography

1. Sparsely distributed conductivity
2. 16 Electrodes
3. No noise ($\delta = 0$)



Electrical Impedance Tomography



THANK YOU FOR YOUR ATTENTION!