NASCA 2018
Numerical Analysis and Scientific Computation with Applications

July, 2–6, 2018, Kalamata, Greece

Acceleration of iterative regularization methods by delta-convex functionals in Banach spaces

Claudio Estatico

Joint work with: Paola Brianzi, Fabio Di Benedetto, Luca Surace.

Department of Mathematics, University of Genoa, Italy

Outline

I - Inverse Problems, deblurring, and Tikhonov-type variational approaches.

II - Minimization of the residual by gradient-type iterative methods in (Hilbert and) Banach spaces.

III - Acceleration of the convergence via operator-dependent penalty terms: the "ir-regularization".

IV - The vector space of the DC (difference of convex) functions, and some relations with linear algebra.

V - Adaptive regularization into the Variable Exponent Lebesgue spaces.

VI - Numerical results in geoscience and in image deblurring (special thanks for the latter to Master student Brigida Bonino).

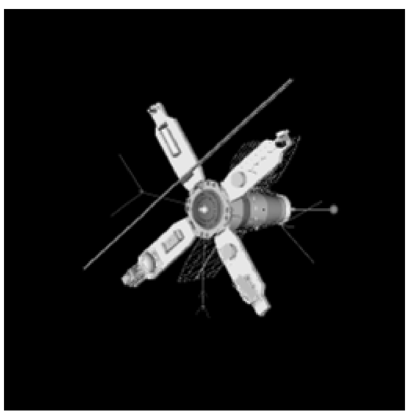
Inverse Problem

By the knowledge of some "observed" data g (i.e., the effect),
find an approximation of some model parameters f (i.e., the cause).

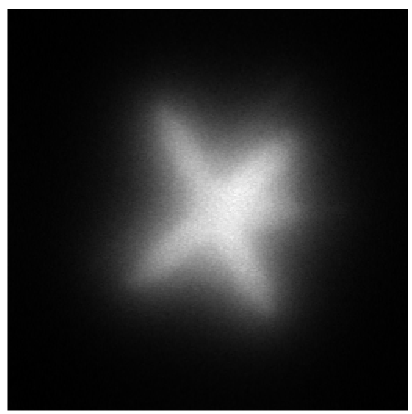
Given the (noisy) data $g \in G$,
find (an approximation of) the unknown $f \in F$ such that

$$Af = g$$

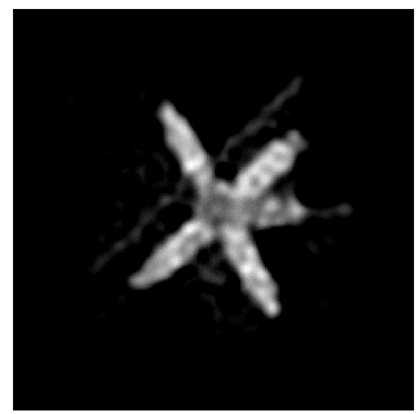
where $A : F \longrightarrow G$ is a known linear operator,
and F,G are two functional (here Hilbert or Banach) spaces.



True image



Blurred (noisy) image



Restored image

Inverse problems are usually ill-posed, they need regularization techniques.

Solution of inverse problems by minimization

Variational approaches are very useful to solve the functional equation

$$Af = g\,.$$

These methods minimize the Tikhonov-type variational functional Φ_α

$$\Phi_\alpha(f) = \frac{1}{p}\|Af - g\|_G^p + \alpha\mathcal{R}(f)\,,$$

where $1 < p < +\infty$, $\mathcal{R} : F \longrightarrow [0, +\infty)$ is a (convex and proper) functional, and $\alpha > 0$ is the regularization parameter.

The "data-fitting" term $\|Af - g\|_G^p$ is called residual (usually in mathematics) or cost function (usually in engineering).

The "penalty" term $\mathcal{R}(f)$ is often $\|f\|_F^q$, or $\|\nabla f\|_F^q$ or $\|Lf\|_F^q$, for $q \geq 1$ (such as the Hölder conjugate of p) and a differential operator L which measures the "non-regularity" of f.

Several regularization methods for ill-posed functional equations by variational approaches have been first formulated as minimization problems in Hilbert spaces (i.e., the classical approach). Later they have been extended to Banach spaces setting (i.e., the more recent approach).

Convex optimization in Banach spaces (such as L^1 for sparse recovery or L^p, $1 < p < 2$ for edge restoration) helps to derive new algorithms.

	Hilbert spaces	Banach spaces
Benefits	Easier computation (Spectral theory, eigencomponents)	Better restoration of the discontinuities; Sparse solutions
Drawbacks	Over-smoothness (bad localization of edges)	Theoretical involving (Convex analysis required)

Minimization of the residual by gradient-type iterative methods

For the Tikhonov-type functional $\Phi_\alpha(f) = \frac{1}{p}\|Af - g\|_G^p + \alpha\mathcal{R}(f)$, the basic minimization approach is the gradient-type iteration, which reads as

$$f_{k+1} = f_k - \tau_k \psi_\alpha(f_k, g)$$

where

$$\psi_\alpha(f_k, g) \approx \partial\left(\frac{1}{p}\|Af - g\|_G^p + \alpha\mathcal{R}(f)\right),$$

i.e., $\psi_\alpha(f_k, g)$ is an approximation of the (sub-)gradient of the minimization functional Φ_α at point f_k, and $\tau_k > 0$ is the step length.

For $\Phi_\alpha(f) = \frac{1}{2}\|Af - g\|_2^2 + \alpha\frac{1}{2}\|f\|_2^2$ in L^2 Hilbert space, from

$$\partial\Phi_\alpha(f) = \nabla\Phi_\alpha(f) = A^*(Af - g) + \alpha f$$

we have the simplest iterative method, i.e., the (mod.) Landweber method,

$$f_{k+1} = f_k - \tau(A^*(Af_k - g) + \alpha f_k)$$

where $\tau \in (0, 2(\|A\|_2^2 + \alpha)^{-1})$ is a fixed step length [Scherzer, 1998].

Minimization in Banach spaces

The computation of the (sub-)differential of the Tikhonov functional requires the (sub-)differential of the norm of the Banach space involved. Here the key tool is the so-called "duality map".

A duality map is a special function which allows us to associate an element of a Banach space B with an element (or a subset of elements) of its dual B^*.

Theorem (Asplund)
Let B be a Banach space and let $p > 1$. A duality map J_B is the subdifferential of the convex functional $f : B \longrightarrow \mathbb{R}$ defined as $f(b) = \frac{1}{p}\|b\|_B^p$:

$$J_B = \partial f = \partial \left(\frac{1}{p} \| \cdot \|_B^p \right) .$$

The (sub-)differential of the residual $\frac{1}{p}\|Af - g\|_G^p$, by means of the duality map J_G , is the following

$$\partial \left(\frac{1}{p} \|Af - g\|_G^p \right) = A^* J_G(Af - g) .$$

Landweber iterative method in Hilbert spaces

$A : F \longrightarrow G \qquad A^* : G \longrightarrow F \qquad \Phi_\alpha(f) = \frac{1}{2}\|Af - g\|_G^2 + \alpha\frac{1}{2}\|f\|_F^2$

$$f_{k+1} = f_k - \tau(A^*(Af_k - g) + \alpha f_k)$$

Landweber iterative method in Banach spaces

$A : F \longrightarrow G \qquad A^* : G^* \longrightarrow F^* \qquad \Phi_\alpha(f) = \frac{1}{p}\|Af - g\|_G^p + \alpha\frac{1}{p}\|f\|_F^p$

$$f_{k+1} = J_{F^*}\Big(J_F f_k - \tau_k(A^* J_G(Af_k - g) + \alpha J_F f_k)\Big)$$

Some remarks.

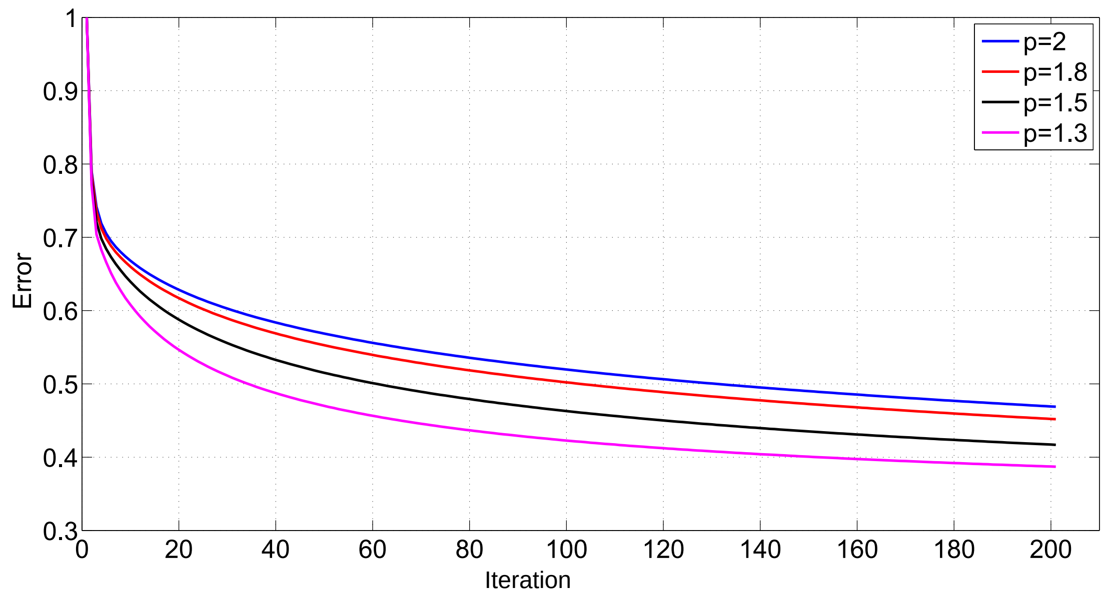
In the Banach space L^p, with $1 < p < +\infty$, we have

$$J_{L^p}(f) = |f|^{p-1}\,\mathrm{sgn}(f)\,.$$

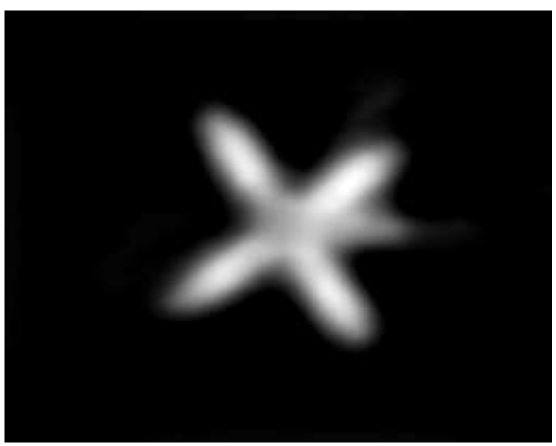
J_{L^p} is a non-linear, single-valued, diagonal operator, which cost $O(n)$ operations, and does not increase the global numerical complexity $O(n\log n)$ of shift-invariant image restoration problems solved by FFT.

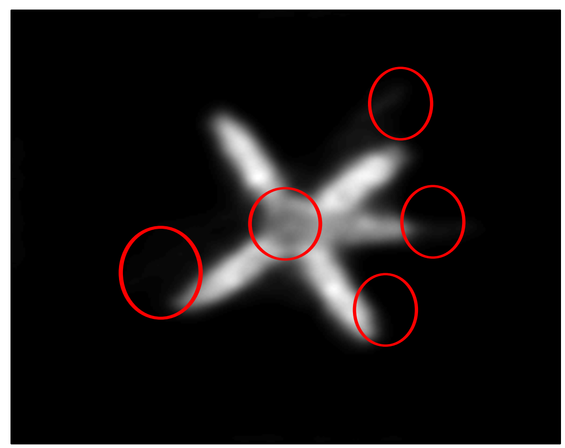
Basically, different L^p spaces leads to different kinds of regularization.

A numerical evidence: L^2 Hilbert vs. L^p Banach spaces



Relative Restoration Errors $RRE(k) = \|f_k - f\|_2 / \|f\|_2$ vs. Iteration Number





Landweber in Hilbert spaces ($p = 2$) Landweber in Banach spaces ($p = 1.3$)

(200-th iteration, with $\alpha = 0$)

Improvement of regularization effects via operator-dependent penalty terms (I)

In the Tikhonov regularization functional $\Phi_\alpha(f) = \frac{1}{p}\|Af - g\|_G^p + \alpha\mathcal{R}(f)$, widely used penalty terms $\mathcal{R}(f)$ include:

(i) $\|f\|^p$, or $\|f - f_0\|^p$, where f_0 is a priori guess for the true solution, whit L^p-norm, $1 < p < +\infty$, or the Sobolev spaces $W^{l,p}$-norm;

(ii) $\|f\|_S^2 = (Sf, f)$ in the Hilbertian case, where $S : F \to F$ is a fixed linear positive definite (often is the Laplacian, $S = \triangle$).

(iii) $\int |\nabla f|$ for Total Variation regularization;

(iv) $\sum_i |(f, \phi_i)|$ or the L^1-norm $\int |f|$ for regularization with sparsity constrains;

In the blue case (ii) with the S-norm, the Landweber iteration becomes:

$$f_{k+1} = f_k - \tau(A^*(Af_k - g) + \alpha S f_k)$$

Improvement of regularization effects via operator-dependent penalty terms (II)

All of the classical penalty terms do not depend on the operator A of the functional equation $Af = g$, but only on f.

On the other hand, it is reasonable that the "regularity" of a solution depends on the properties of the operator A too.

Recalling that, in inverse problems:

Spectrum of A^*A	$\longleftrightarrow$	Subspace	Components
$\lambda(A^*A)$ small	$\longleftrightarrow$	Noise Space	High Frequencies
$\lambda(A^*A)$ large	$\longleftrightarrow$	Signal Space	Low Frequencies

The idea: [T. Huckle and M. Sedlacek, 2012]
The penalty term should measure "how much" the solution f is in the noise space, which depends on A.

Improvement of regularization effects via operator-dependent penalty terms (III)

In [HS12], the key proposal is based on the following operator S

$$S = \left(I - \frac{A^*A}{\|A\|^2} \right),$$

so that $\|f\|_S^2 = (Sf, f) \approx \begin{cases} \textit{large} & \text{if } f \text{ is heavily in the noise space of } A \\ \textit{small} & \text{if } f \text{ is lightly in the noise space of } A \end{cases}$

The linear (semi-)definite operator S is a high pass-filter (please notice that S is NOT a regularization filter, which are all low pass...).
This way, the S-norm is able to measure the "non-regularity" w.r.t. the properties of the actual model-operator A.

In the previous literature, the Tikhonov regularization functional $\Phi_\alpha(f) = \|Af - g\|^2 + \alpha\|f\|_S^2$ is solved, only in Hilbert spaces, by direct methods via Euler-Lagrange normal equations. This way, the direct solver benefits of the very high regularization effects given by $\|f\|_S^2$.

Minimization of Tikhonov functional with high-pass filter S as penalty term

We can apply the idea of the high-pass filter S to any iterative method in Banach spaces for the minimization of the Tikhonov regularization functional

$$\Phi_\alpha(f) = \|Af - g\|_G^p + \alpha\|f\|_S^p \,.$$

Drawback: The high-regularization effects of the S-norm slow down the convergence of the iterations.

The operator S reduces the components in the signal space and keep the component in the noise space. But we have to do the opposite. So we modify the (convex) Tikhonov functional into the non-convex (family of) functionals

$$\tilde{\Phi}_{\alpha_k}(f) = \|Af - g\|_G^p - \alpha_k\|f\|_S^p \,.$$

This way, the basic Landweber iterations becomes

$$f_{k+1} = J_{F^*}\Big(J_F f_k - \tau_k(A^* J_G(Af_k - g) - \alpha_k S J_F f_k)\Big) \,,$$

The action of $-S$ is a *"ir-regularization"*, which reduces the (over-smoothing) regularization effects of the iterative method (in the first iterations ...).

Minimization of difference of convex functions

Any *ir-regularization* functional

$$\tilde{\Phi}_{\alpha_k}(f) = \|Af - g\|_G^p - \alpha_k \|f\|_S^p \,.$$

is not convex. However, it is composed by the difference of two convex functional, $\|Af - g\|_G^p$ and $\|f\|_S^p$.

These kind of functions, called DC -difference of convex- functions (or delta-convex functions), have been exhaustively analyzed since about 1950.

The class of DC functions is a remarkable subclass of locally Lipschitz functions that is of interest both in analysis and optimization.

It is naturally the smallest vector space containing all continuous convex functions on a given set. And it is surprisingly "large"!

If you know the DC decomposition (as in our case), there exist algorithms for the global minimum based on primality-duality.

The vector space of DC (difference of convex) functions

Let DC(X) be the vector space of scalar DC functions on the open convex set $X \subset \mathbb{R}^n$ (Euclidean case). Let

$$h = h_1 - h_2 \in \mathrm{DC}(X)$$

where both h_1 and h_2 are convex functions on X. Clearly h_1 and h_2 can be chosen non-negative.

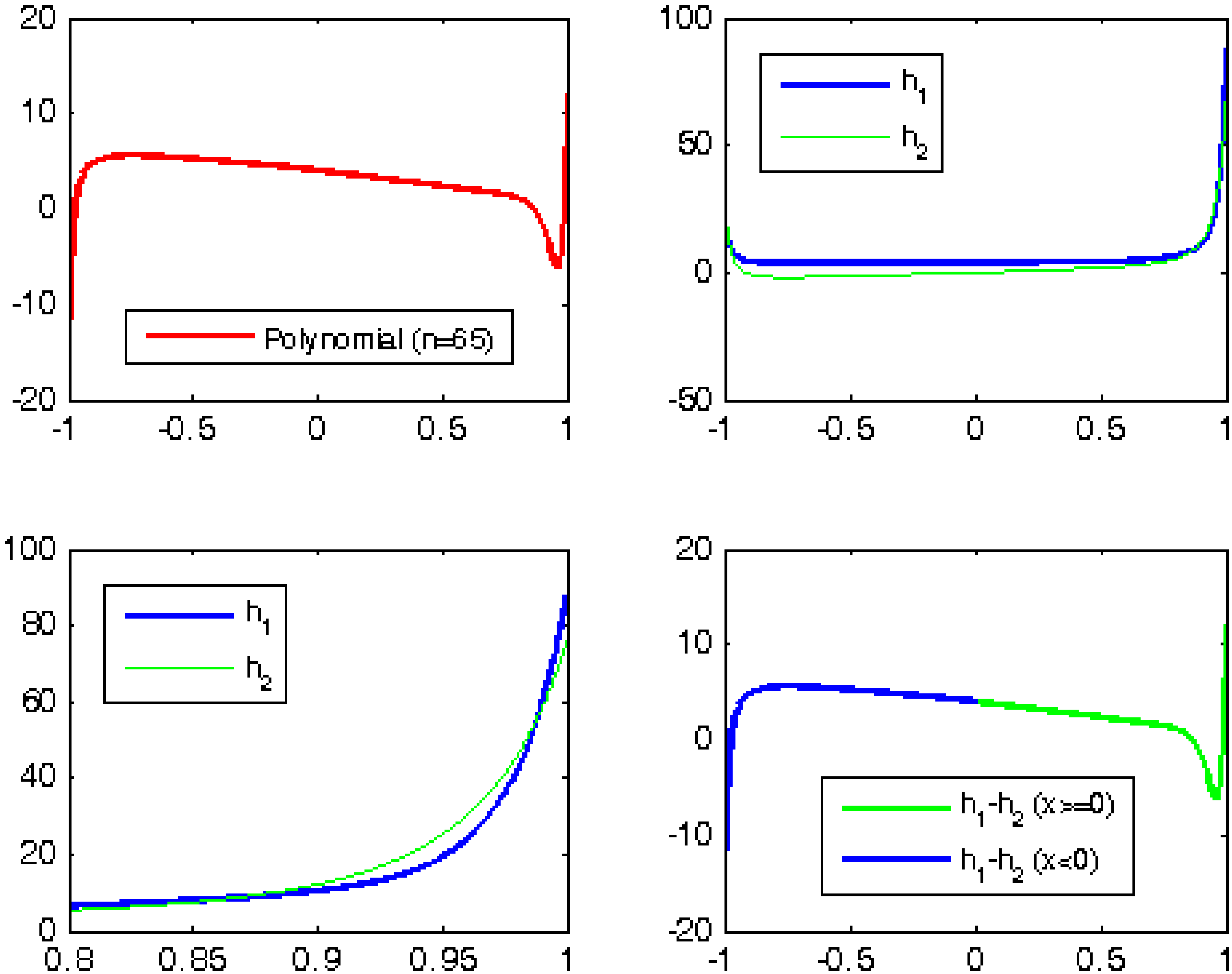
The class of DC functions is a subclass of locally Lipschitz functions.

(i) In the simplest case $n = 1$, $h(x)$ is a DC function if and only if it has left and right derivatives and these derivatives are of bounded variation on every closed bounded interval interior to X (that is, f' is a difference of two nondecreasing functions).

(ii) Thanks to (i), any polynomial on $\mathbb{R}^n$ is DC. Indeed, each polynomial p can be decomposed as $p = q - r$ where r, q are nonnegative convex functions. Easy proof: $x^{2n+1} = (x^+)^{2n+1} - (x^-)^{2n+1}$ and x^{2n} are DC

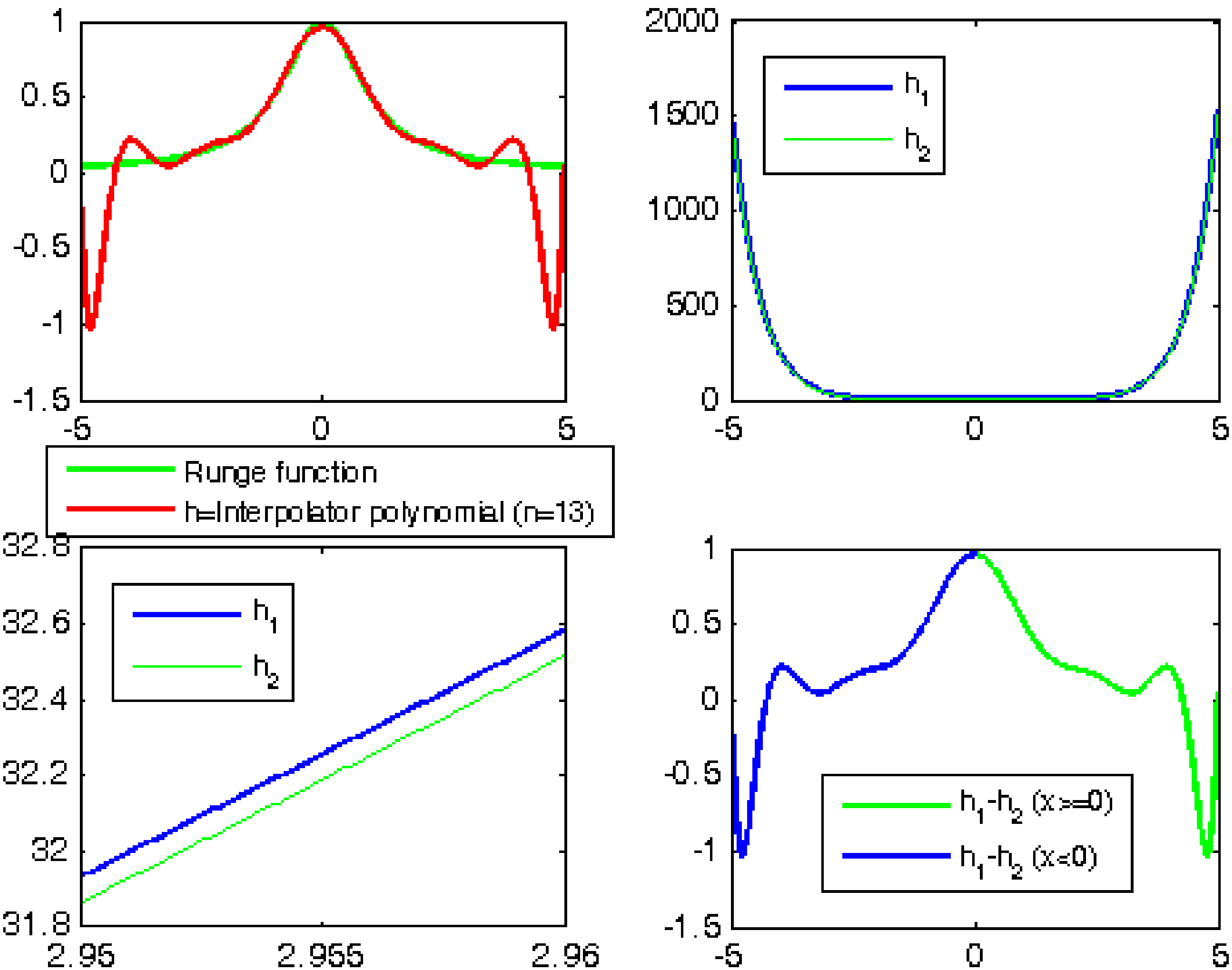
(iii) Thanks to (ii), DC functions are dense uniformly in $C(X)$ for compact X. Any continuous function can be well approximated by DC functions.

DC decomposition of a (random) polynomial



$$h = h_1 - h_2$$

DC decomp. of the interp. polynomial of Runge function



$$h = h_1 - h_2$$

A sketch on linear algebra: Eigenvalues and DC functions (I)

Let A be a $n \times n$ symmetric matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.

(i) The quadratic form

$$R(x) = \frac{1}{2} x^t A x$$

is DC on $\mathbb{R}^n$.

Indeed, there are many decomposition with positive semi-definite A^+ and A^- such that $R(x) = \frac{1}{2} x^t A^+ x - \frac{1}{2} x^t A^- x$.

(ii) The kth-largest eigenvalue function

$$\lambda_k : A \to \lambda_k(A)$$

is DC on the space of symmetric matrices.

Proof: $\lambda_k = \sum_{j=1}^{j=k} \lambda_j - \sum_{j=1}^{j=k-1} \lambda_j$, and the sum of the first largest eigenvalues is convex, i.e.,

$\sum_{j=1}^{j=k} \lambda_j(tA_1 + (1-t)A_2) \leq t \sum_{j=1}^{j=k} \lambda_j(A_1) + (1-t) \sum_{j=1}^{j=k} \lambda_j(A_2)$.

Eigenvalues and DC functions (II)

Let A be a $n \times n$ symmetric positive definite matrix with eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$.

There are some further more involving facts. An example:

From the Rayleigh quotient, we know that the largest eigenvalue is

$$\lambda_1 = \max\{x^t A x \ : \ \|x\| \leq 1\}\,.$$

Via dualization schemes for convex constrains, the optimization theory for DC function shows that

$$\lambda_1 = -\min\{x^t A^{-1} x - 2\|x\| \ : \ x \in \mathbb{R}^n\}\,, \qquad \text{as well as}$$

$$\lambda_1 = -\min\{\|x\|^2 - 2\sqrt{x^t A x} \ : \ x \in \mathbb{R}^n\}\,.$$

Eigenvalues and DC functions (III)

We verify the first one (the second one is similar). Recall that A is a $n \times n$ symmetric positive definite matrix with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n > 0$.

$$\lambda_1 = -\min\{x^t A^{-1} x - 2\|x\| \,:\, x \in \mathbb{R}^n\}$$

Indeed, let u be fixed, with $\|u\| = 1$, and consider $x = \varrho u$, with $\varrho \geq 0$.

For $x^t A^{-1} x - 2\|x\| = (u^t A^{-1} u)\varrho^2 - 2\varrho := s(\varrho)$, the minimum is attained for $s'(\tilde{\varrho}) = 2(u^t A^{-1} u)\tilde{\varrho} - 2 = 0$, that is, at $\tilde{\varrho} = (u^t A^{-1} u)^{-1}$.

This minimum is:
$s(\tilde{\varrho}) = (u^t A^{-1} u)(u^t A^{-1} u)^{-2} - 2(u^t A^{-1} u)^{-1} = -(u^t A^{-1} u)^{-1}$.

Searching for the minimum of all the minima of $s(\tilde{\varrho})$ (i.e, all over the directions u), we obtain:

$\min\{x^t A^{-1} x - 2\|x\| \,:\, x \in \mathbb{R}^n\} = \min_{\|u\|=1}\{-(u^t A^{-1} u)^{-1}\} =$
$= -\max_{\|u\|=1}\{(u^t A^{-1} u)^{-1}\} = -(\lambda_{\min}(A^{-1}))^{-1} = -\lambda_{\max}(A) = -\lambda_1$.

The ir-regularization method (in Banach spaces)

We consider the Landweber method in Banach spaces, with ir-regularization penalty term (i.e., minim. of a DC func.). We write the iteration as follows:

$$f_{k+1} = J_{F^*}\left(J_F f_k - \tau A^* J_G(Af_k - g) + \beta_k S J_F f_k\right),$$

$$S = \left(I - \frac{A^* J_G A}{\|A\|\|A^*\|}\right),$$

Theorem 1 (Noiseless case) *[Brianzi, Di Benedetto, E., Surace, 2018]*
If $(\beta_k)_k$ *is positive and decreasing, with* $\sum_{k=0}^{+\infty} \beta_k < +\infty$, *and* $\tau \in \left(0, \frac{2}{\|A\|^2}\right)$, *then*

$$\lim_{k\to\infty} f_k = f^\dagger,$$

for the ir-regularization iterative method in Hilbert spaces.

The ir-regularization method for noisy data

Via filter factor analysis, we can prove that acceleration via ir-regularization still converges and gives rise to a regularization method. Recall that, by singular-value decomposition $(\sigma_n; v_n, u_n)_{n=1}^{+\infty}$ of compact operators

$$y = Ax = \sum_{n=1}^{\infty} \sigma_n (x, v_n) u_n \qquad \text{and} \qquad x^\dagger = A^\dagger y = \sum_{n=1}^{\infty} \frac{(y, u_n)}{\sigma_n} v_n \ ,$$

where $\sigma_n > 0 \ \ \forall s$ and $\sigma_n \to 0$ for $n \to +\infty$ $(\sigma_1 \geq \sigma_2 \geq \dots)$.

We write the k-th iteration x_k as filtered version of $x^\dagger$:

$$x_k = \sum_{n=1}^{\infty} \phi_k(\sigma_n) \frac{(y, u_n) v_n}{\sigma_n} \ .$$

By classical reegularization theory in Hilbert space,
if $\forall \sigma \in (0, \|A\|]$, $\lim_{k \to +\infty} \phi_k(\sigma) = 1$, and same stability conditions hold (...), the iterative scheme is a regularization algorithm.

Theorem 2 (Noisy case) *The ir-regularization method is a regularization algorithm.*

The acceleration given by the ir-regularization method (I)

By manipulation of recursions of both the basic and the accelerated methods, two derive the following formulas:

$$1 - \phi_{k+1}^{\text{Land}}(\sigma) = (1 - \tau\sigma^2)(1 - \phi_k^{\text{Land}}(\sigma)) \; ,$$

$$1 - \phi_{k+1}^{\text{irr}}(\sigma) = (1 - \tau\sigma^2)(1 - \phi_k^{\text{irr}}(\sigma)) + \beta_k \left(1 - \frac{\sigma^2}{\sigma_1^2}\right) \phi_k^{\text{irr}}(\sigma) \; .$$

Assume now that, for a certain n, the inequalities $0 < 1 - \phi_k^{\text{irr}}(\sigma) \leq$ $\leq 1 - \phi_k^{\text{Land}}(\sigma) < 1$ hold (this is trivially true for $n = 0$). We obtain

$$1 - \phi_{k+1}^{\text{irr}}(\sigma) \leq 1 - \phi_{k+1}^{\text{Land}}(\sigma) + \beta_k \left(1 - \frac{\sigma^2}{\sigma_1^2}\right) \phi_k^{\text{irr}}(\sigma) \; .$$

We see from the formula that the new ir-regularization filter factor can approach the limit 1 faster than the corresponding of classical Landweber.

The acceleration given by the ir-regularization method (II)

The first values of the filter factors for the classical Landweber and the accelerated by ir-regularization for a constant ir-regularization parameter ($\beta_k \equiv 0.1; \tau = 0.8, \sigma_1 = 1$).

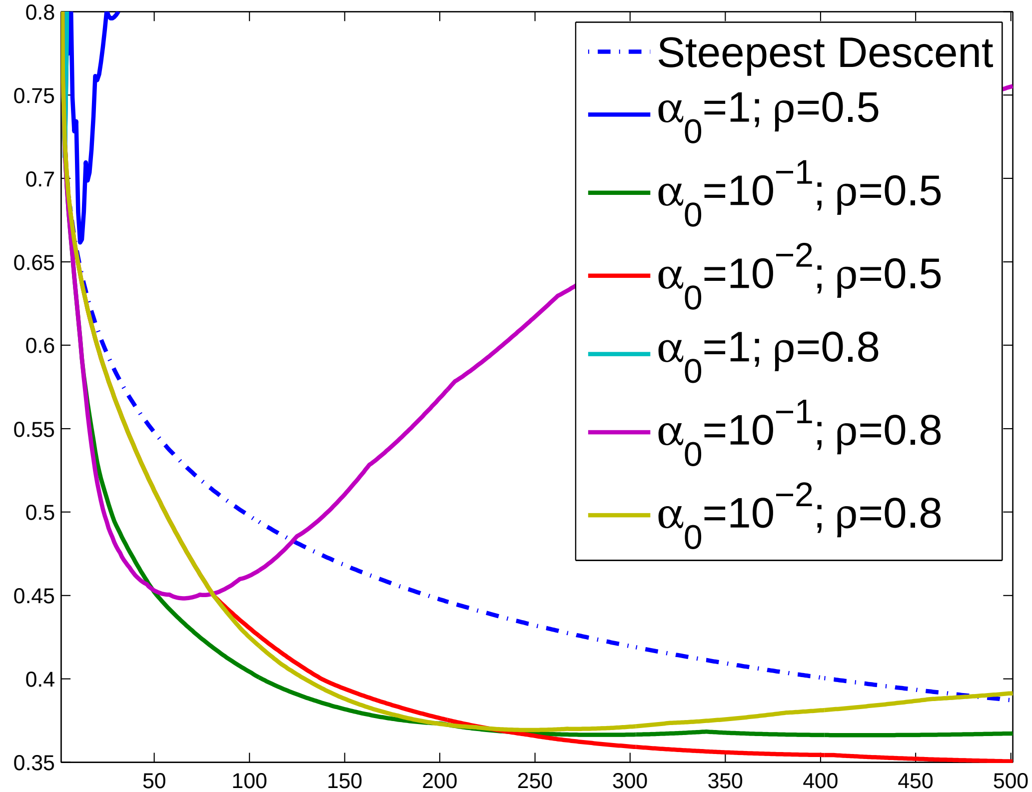
k	$\phi_k^{\text{Land}}(\sigma)$	$\phi_k^{\text{irr}}(\sigma)$
1	0.6400	0.6400
2	0.8704	0.8934
3	0.9533	0.9938
4	0.9832	1.0335
5	0.9940	1.0493
6	0.9978	1.0555

The values of the ir-regularization method approach the value 1 (which means "no filtering") faster than the Landweber one.

Notice that, in practice, the acceleration parameter β_k must be reduced in order to prevent that $\phi_k^{\text{irr}}(\sigma)$ has values larger than 1.

Numerical results (I) (Satellite data set)

Relative Restoration Errors $RRE(k) = \|f_k - f\|_2 / \|f\|_2$ vs. Iteration Number



$\beta_k = \varrho^{s_k}\alpha_0$ (ir-regularization decreasing sequence); $\quad F = G = L^2$.

Adaptive update of the ir-regularization sequence (β_k):

With $s_0 = 0$, **if** $\|Af_k - g\|_2 \geq \|Af_{k-1} - g\|_2$ **then** $s_k = s_{k-1} + 1$,
else $s_k = s_{k-1}$.

Numerical results (II): The preconditioned version

Next, we consider the preconditioned version of the method, where the preconditioner D is a regularization preconditioner in the dual space, that is, $D : G^* \longrightarrow G^*$.

The dual-preconditioner D is built by means of (an extension to Banach spaces of) a filtering procedure of the T.Chan preconditioner in Hilbert space.

$$f_{k+1} = J_{F^*}\Big(J_F f_k - \tau_k D A^* J_G(Af_k - g) + \beta_k S J_F f_k\Big) ,$$

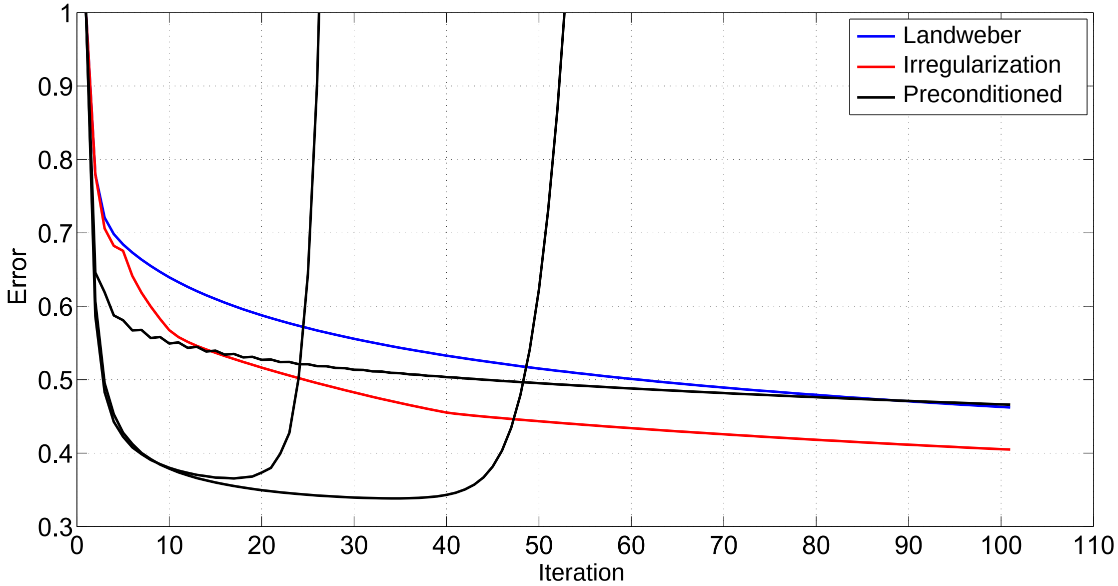
$$S = \left(I - \frac{A^* J_G A}{\|A\| \|A^*\|}\right) .$$

The preconditioners allow to speed up the convergence, but usually they give rise to instability (that is, to a fast semi-convergence).

The classical preconditioned method is faster than the method with ir-regularization. However, surprisingly enough, the ir-regularization improves the stability of the preconditioned method.

Numerical results (II): preconditioner VS ir-regularization

Relative Restoration Errors $RRE(k) = \|f_k - f\|_2 / \|f\|_2$ vs. Iteration Number

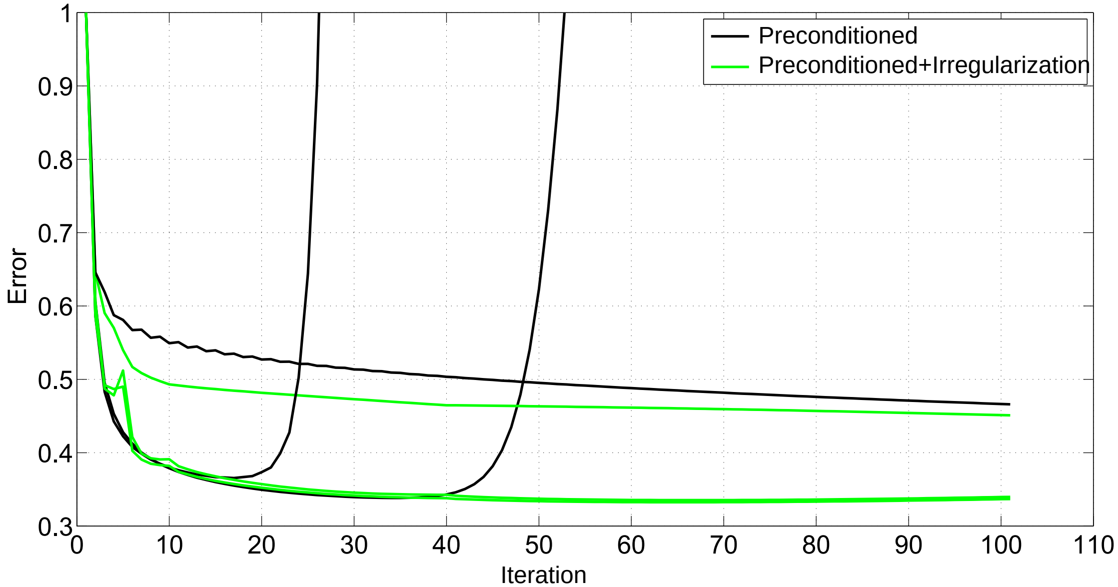


ir-regularization at 100-th iteration

preconditioner at 36-th iteration.

Numerical results (II): preconditioner AND ir-regularization

Relative Restoration Errors $RRE(k) = \|f_k - f\|_2 / \|f\|_2$ vs. Iteration Number



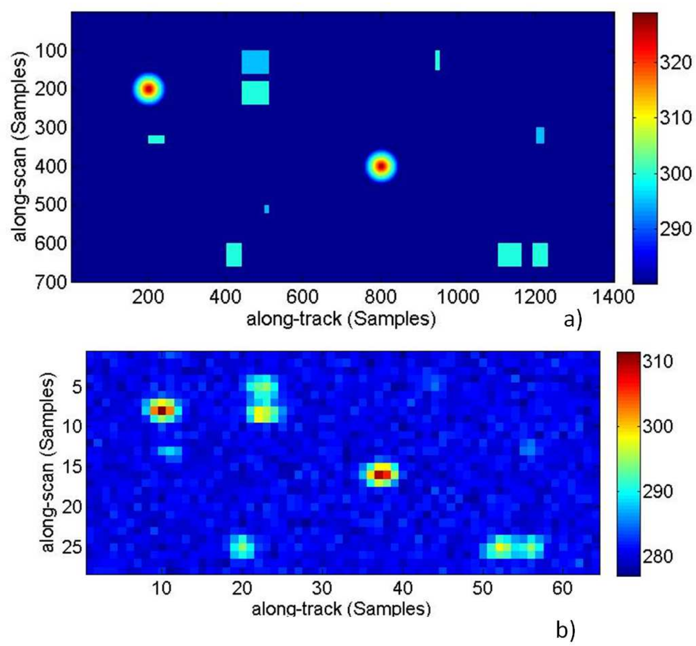
preconditioner, at 36-th iteration

prec. AND ir-regulariz. at 36-th iteration.

Numerical results (III): A geophysical application (Hilbert)

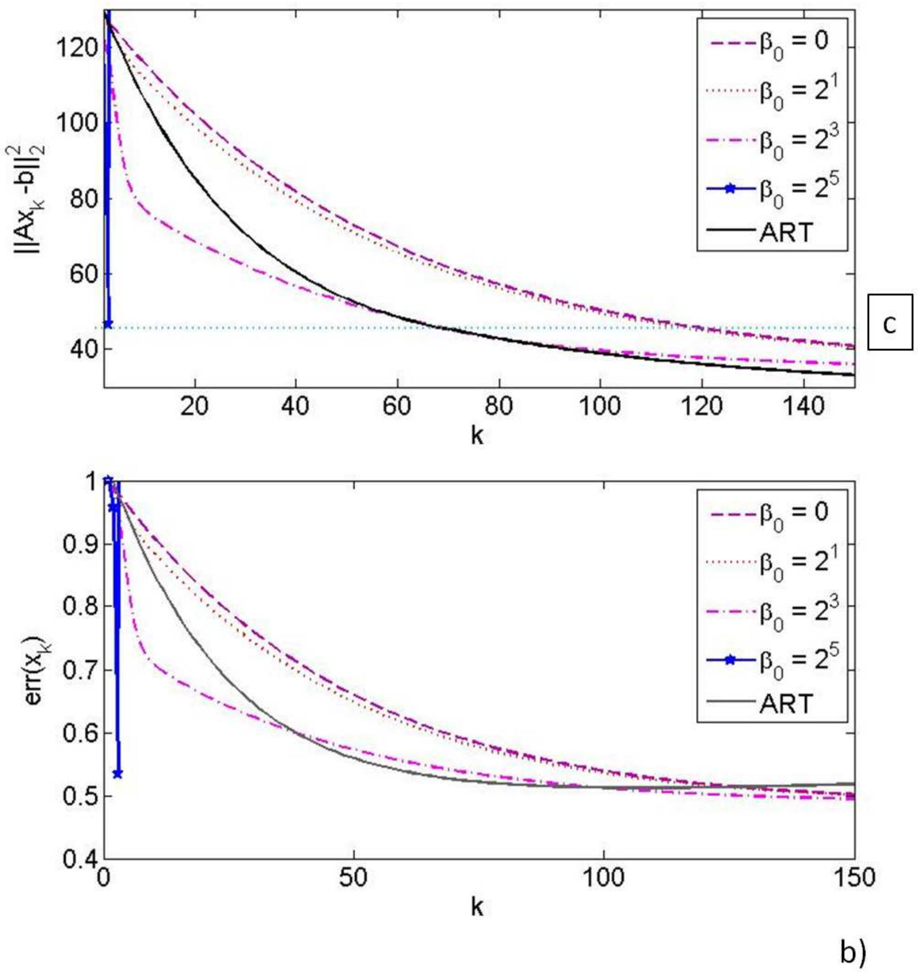
Aim: To enhance the spatial resolution of simulated Special Sensor Microwave/Imager (SSM/I) radiometer measurements (in Hilbert spaces).

Unknown: brightness temperature on a 1400 × 700 km Earth's surface
Data: remotely sensed measurements via Fredholm integral operator
Noise: 10% Gaussian, zero mean (Courtesy: monde-geospatial.com)



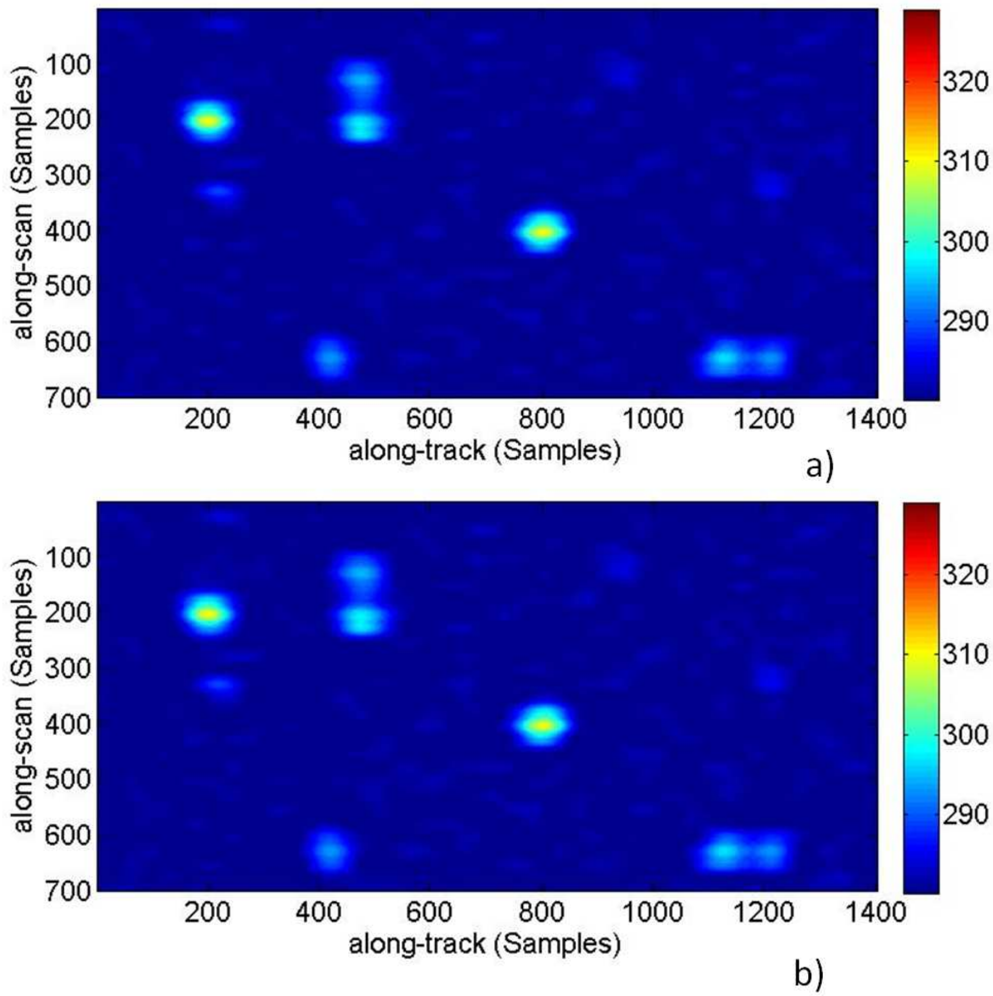
(a) Reference field - (b) Simulated noisy measurements

Numerical results (III): A geophysical application



TOP: 2-norm of the residual versus the iteration index k.
BOTTOM: Relative error between the k-reconstructed and the reference field versus k. c is an upper bound related to the 2-norm of the noise.
ART = algebraic reconstruction technique (to compare with). $\beta_k = \beta_0/2^k$.

Numerical results (III): A geophysical application



TOP: Reconstructed field using the conventional Landweber method ($\beta_0 = 0\,,\;\; k = 175$).
BOTTOM: Reconstructed field using the improved Landweber method ($\beta_0 = 2^3\,,\;\; k = 120$).
Relative restoration error: 0.53, both.

The new framework: variable exponent Lebesgue spaces

A recent improvement: the ill-posed functional equation $Af = g$ is solved in the $L^{p(\cdot)}$ Banach space, namely, the variable exponent Lebesgue spaces (a special case of the so-called Musielak-Orlicz functional spaces).

In a variable exponent Lebesgue spaces, to measure a function f, instead of a constant exponent p all over the domain, we have a pointwise variable (i.e., a distribution) exponent $1 \leq p(\cdot) \leq +\infty$:

$$\int |f(x)|^{p(x)} dx \,.$$

This way, different regularization levels on different regions of the image to restore can be automatically and adaptively assigned. Different pointwise regularization is useful because background, low intensity, and high intensity values require different filtering levels (see Nagy, Pauca, Plemmons, Torgersen, J Opt Soc Am A, 1997, "Regularization is accomplished by varying the preconditioners -i.e., the regularization levels- across the segments").

The norm of the variable exponent Lebesgue space

In the conventional case L^p, the norm is $\|f\|_{L^p} = \left(\int |f(x)|^p dx\right)^{1/p}$.

In $L^{p(\cdot)}$ Lebesgue spaces, the definition and computation of the norm is not straightforward, since we have not a constant value for computing the (mandatory) radical. How is the following problem solved?

$$\|f\|_{L^{p(\cdot)}} = \left(\int |f(x)|^{p(x)} dx\right)^{1/???}.$$

The solution: compute first the so-called modular (for $1 < p(\cdot) < +\infty$)

$$\varrho_{L^{p(\cdot)}}(f) = \int |f(x)|^{p(x)} dx\,,$$

and then obtain the norm by solving a 1D minimization problem

$$\|f\|_{L^{p(\cdot)}} = \inf\left\{\lambda > 0 : \varrho_{L^{p(\cdot)}}(f/\lambda) \leq 1\right\}.$$

In the case of constant distribution $p(\cdot) = p$, this norm coincides with the classical one $\|f\|_{L^p}$.

The duality map of the variable exponent Lebesgue space

By extending the duality maps, we can use the previous ir-regularization iterative method in the framework of variable exponent Lebesgue spaces.

For any constant $1 < r < +\infty$, we recall that the duality map, that is, the (sub-)differential of the functional $\frac{1}{r}\|f\|_{L^p}^r$, in the classical Banach space L^p, with constant $1 < p < +\infty$, is defined as follows

$$\Big(J_{L^p}(f)\Big)(x) = \frac{|f(x)|^{p-1}\,\mathrm{sgn}(f(x))}{\|f\|_p^{p-r}}\,.$$

By generalizing a result of P. Matei [2012], we have that the corresponding duality map in variable exponent Lebesgue space is defined as follows

$$\Big(J_{L^{p(\cdot)}}(f)\Big)(x) = \frac{1}{\int_\Omega \frac{p(x)\,|f(x)|^{p(x)}}{\|f\|_{p(\cdot)}^{p(x)}}\,dx}\;\frac{p(x)\,|f(x)|^{p(x)-1}\,\mathrm{sgn}(f(x))}{\|f\|_{p(\cdot)}^{p(x)-r}}\,,$$

where any product and any ratio have to be considered as pointwise.

The adaptive algorithm in variable exponent Lebesgue spaces

It is a numerical evidence that, in L^p image deblurring,

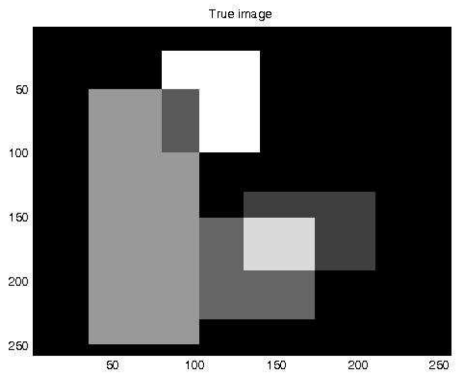
- dealing with small $1 \approx p << 2$ improves sparsity and allows a better restoration of the edges of the images and of the zero-background,
- dealing with $p = 2$ (or $p > 2$), allows a better restoration of the pixels with the highest intensities.

The idea: to use (a re-blurred and scaled into $[1, 2]$ version of) the blurred data g as distribution of the exponent $p(\cdot)$ for the variable exponent Lebesgue spaces $L^{p(\cdot)}$ where computing the solution. Example:

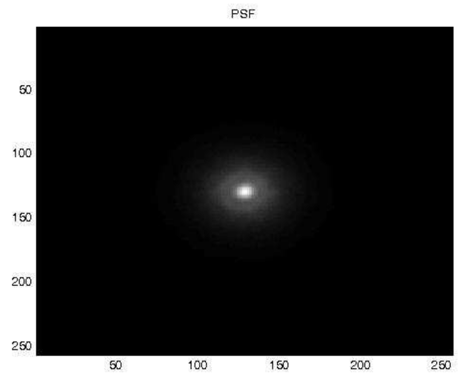
$$p(\cdot) = 1 + [A^*g(\cdot) - \min(A^*g)]/[\max(A^*g) - \min(A^*g)]$$

The Landweber (i.e., fixed point) iterative scheme in this $L^{p(\cdot)}$ Banach space can be modified as adaptive iteration algorithm, by recomputing, after each fixed number of iterations, the exponent distribution $p_k(\cdot)$ by means of the k-th restored image f_k (instead of the first re-blurred data Ag), that is

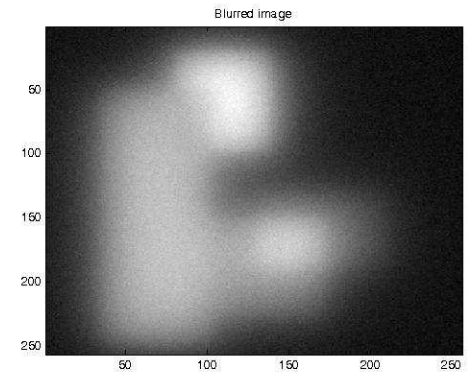
$$p_k(\cdot) = 1 + [f_k(\cdot) - \min(f_k)]/[\max(f_k) - \min(f_k)]$$



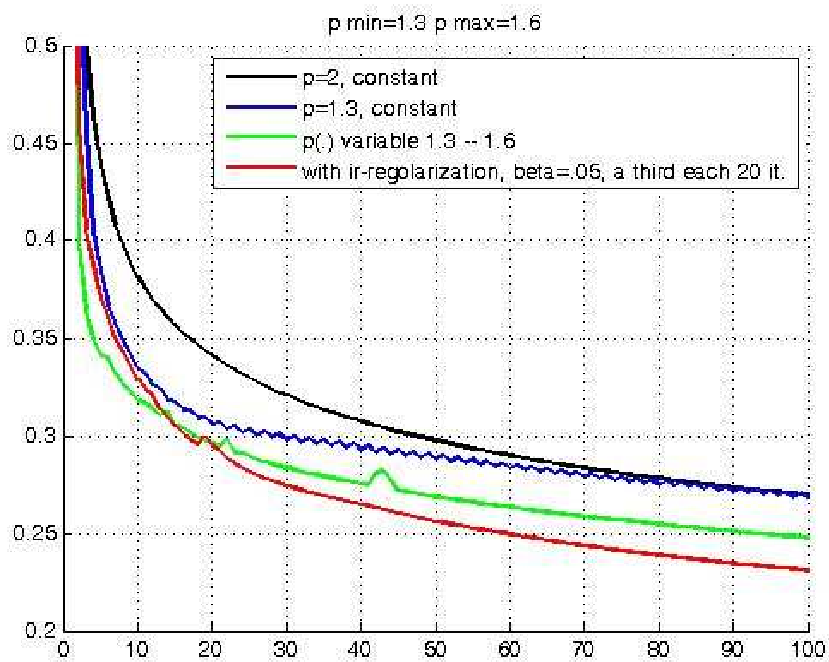
True image

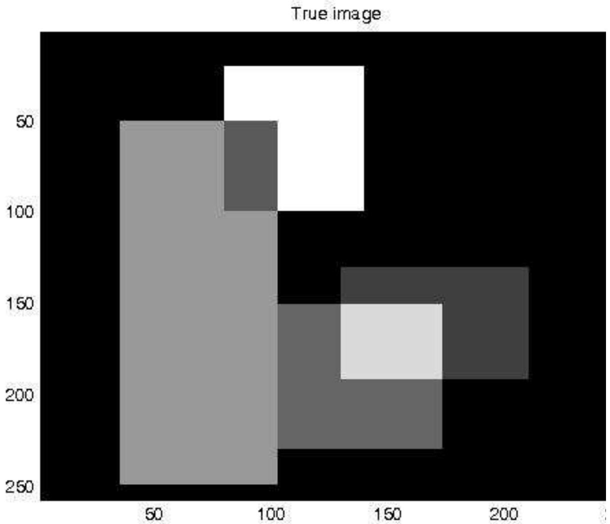


Point Spread Function

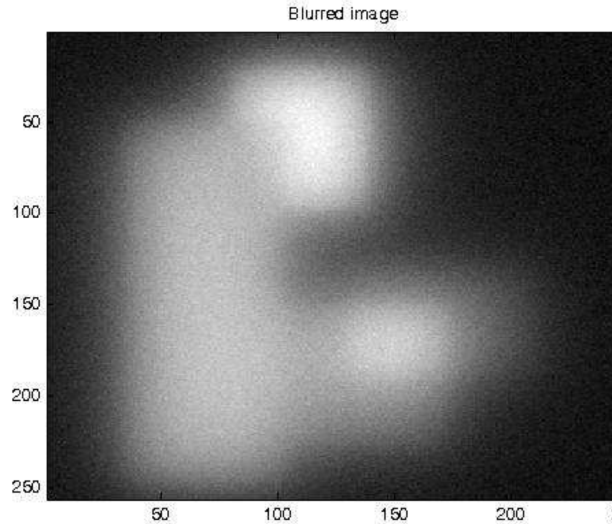


Blurred image (noise = 4.7%)

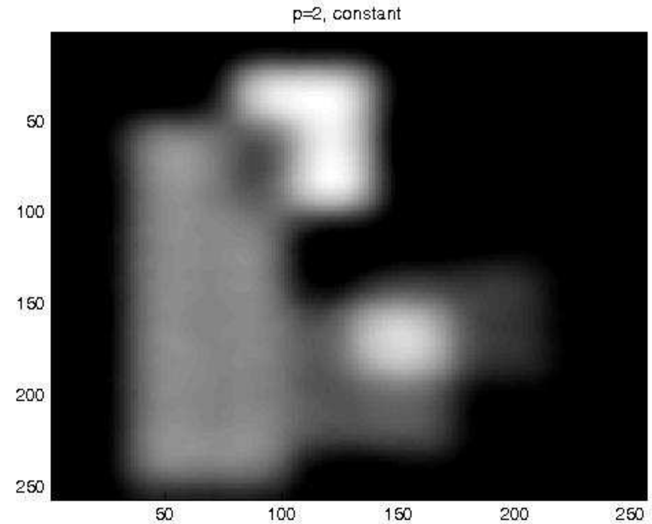




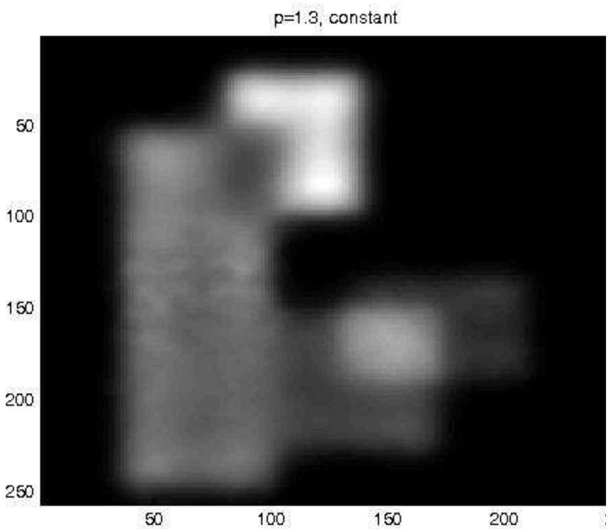
True image



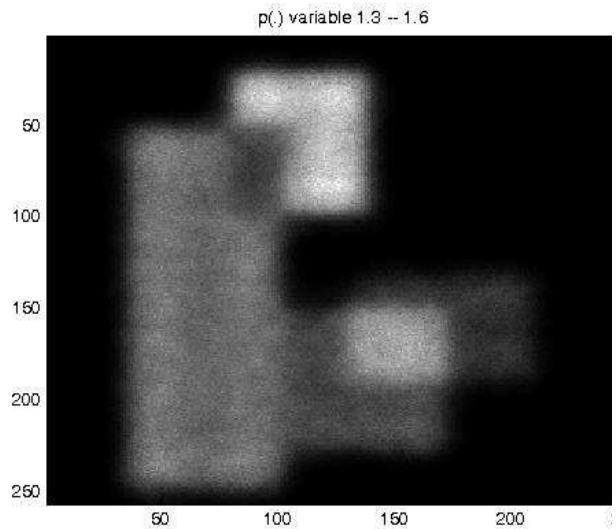
Blurred (noise = 4.7%)



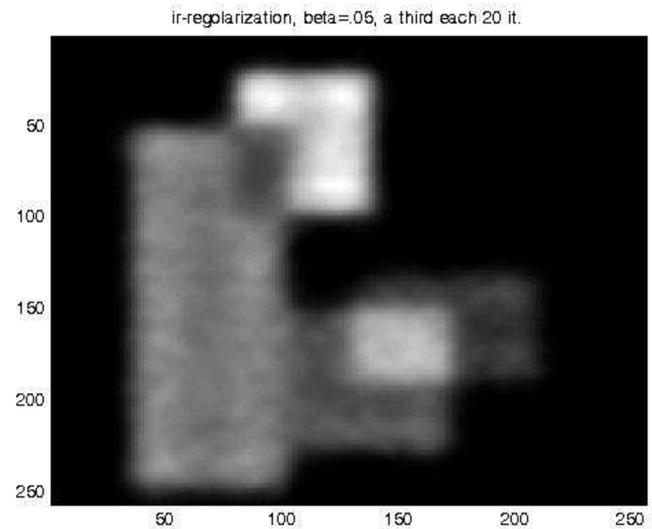
$p = 2 \quad (0.2692)$



$p = 1.3 \quad (0.2681)$



$p = 1.3 - 1.6 \quad (0.2473)$



$1.3 - 1.6$ and irreg. (0.2307)

Conclusions

- Regularization iterative methods can be accelerated by ir-regularization.
- Theory for global minimization of DC functions via dualization could be useful to speed-up iterative schemes.
- Extension of regularization algorithms to variable Lebesgue spaces could provide adaptive regularization too...

Thank you for your attention.

References

[1] T. Huckle, M. Sedlacek,
Tykhonov-Phillips regularization with operator dependent seminorms,
Numer. Alg., vol. 60, pp. 339–353, 2012.

[2] F. Lenti, F. Nunziata, C. Estatico, M. Migliaccio,
Spatial resolution enhancement of Earth observation products using an acceleration technique for iterative methods, *IEEE Geosc. Rem. Sens. Lett.*, vol. 12, pp. 269-273, 2015.

[3] P. Brianzi, F. Di Benedetto, C. Estatico,
Preconditioned iterative regularization in Banach spaces,
Comp. Optim. Appl., vol. 54, pp. 263–282, 2013.

[4] P. Brianzi, F. Di Benedetto, C. Estatico, L. Surace
Irregularization accelerates iterative regularization,
Calcolo, vol. 55:17, June 2018.

[5] Y. Eidelman, V. Milman, A. Tsolomitis,
Functional Analysis: an Introduction.
American Mathematical Society, Providence, R.I., 2004.

[6] L. Diening, P. Harjulehto, P. Hästö, M. Ruzicka,
Lebesgue and Sobolev Spaces with Variable Exponents,
Lecture Notes in Mathematics, vol. 2017, Springer, 2011.