Numerical Linear Algebra and Applications (NL2A)

24–24 October 2016, CIRM, Luminy, Marseille

Iterative regularization in variable exponent Lebesgue spaces

Claudio Estatico [1]

Joint work with: Brigida Bonino [1], Fabio Di Benedetto [1], Flavia Lenti [2]

[1] Dipartimento di Matematica, Università di Genova, Italy

[2] EUMETSAT, Darmstadt, Germany
(ex ENSEEIHT, Université de Toulouse, France)

Outline

I - Inverse Problems, image deblurring, and regularization by iterative methods in (Hilbert and) Banach spaces.

II - A sketch on variable exponent Lebesgue spaces.

III - Application of variable exponent Lebesgue spaces in iterative regularization methods.

IV - Numerical results in image deblurring.

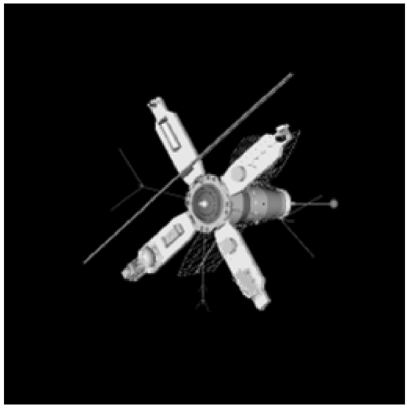
Inverse Problem

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

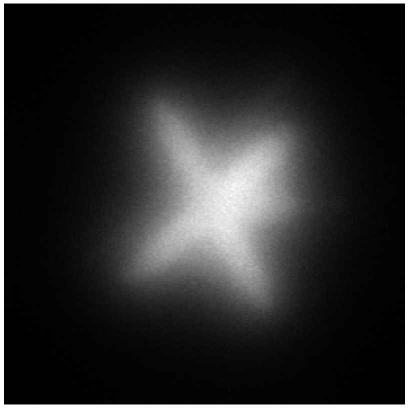
Given the (noisy) data $g \in G$,
find (an approximation of) the unknown $f \in F$ such that

$$Af = g$$

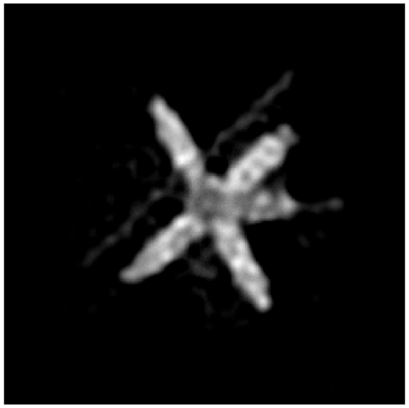
where $A : F \longrightarrow G$ is a known linear operator,
and F,G are two functional (here Hilbert or Banach) spaces.



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

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

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).

Examples: L^1 for sparse recovery or L^p, $1 < p < 2$ for edge restoration.

	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 simple residual functional $\Phi_\alpha(f) = \|Af - g\|_G^p$, 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

$$\psi_\alpha(f_k, g) \approx \partial \left(\|Af - g\|_G^p \right),$$

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

For the least square functional $\Phi_\alpha(f) = \frac{1}{2}\|Af - g\|_2^2$ in L^2 Hilbert space, since

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

we have the simplest iterative method, i.e., the Landweber method,

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

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

From Hilbert to Banach spaces

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

Formally, A^* is the dual operator of A, that is, the operator

$$A^* : G^* \longrightarrow F^* \qquad \text{such that}$$

$$g^*(Af) = (A^*g^*)(f)\,, \quad \forall f \in F \;\text{ and }\; \forall g^* \in G^*\,,$$

where F^* and G^* are the dual spaces of F and G.

If F and G are Hilbert spaces, then F is isometrically isomorph to F^* and G is isometrically isomorph to G^* (by virtue of Riesz Theorem), and the operator $A^* : G^* \longrightarrow F^*$ can be identified with $A^* : G \longrightarrow F$.

However, in general Banach spaces are not isometrically isomorph to their duals. This way, the Landweber iteration above is well defined in Hilbert spaces (...only!)

The key point: To generalize from Hilbert to Banach spaces we have to consider the so-called duality maps.

Minimization in Banach spaces by duality maps

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

Theorem (Asplund [1968])
Let B be a Banach space and $p > 1$. A duality map $J_B : B \longrightarrow 2^{B^*}$ is the sub-differential of the convex functional f defined as $f(b) = \frac{1}{p}\|b\|_B^p$:

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

By chaining rule, the (sub-)differential of the residual $\frac{1}{p}\|Af - g\|_G^p$, by means of the duality map $J_G : G \longrightarrow 2^{G^*}$, is the following

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

Landweber iterative method in Hilbert spaces

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

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

Landweber iterative method in Banach spaces

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

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

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.

Landweber iterative method in Hilbert spaces

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

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

Landweber iterative method in Banach spaces

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

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

Some remarks.

Numerical linear algebra –developed usually in Hilbert spaces– is still useful in Banach spaces (preconditioning, trigonometric matrix algebras for boundary conditions, . . .). It is “inside” all our discretized setting.

Duality maps are the basic tool for generalizing classical iterative methods for linear systems to Banach spaces: Landweber method, CG, Mann iterations, Gauss-Newton Gradient type iterations (for nonlinear problems).

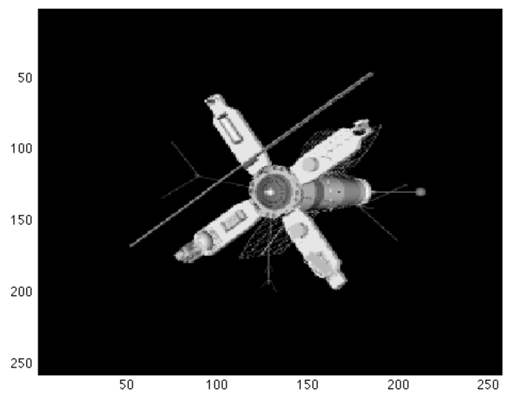
Basic hypotheses for the convergence (and regularization behavior): uniformly smooth and (uniformly) convex Banach spaces. [Schöpfer, Louis, Hein, Scherzer, Schuster, Kazimierski, Kaltenbacher, Q. Jin, Tautenhahn, Neubauer, Hofmann, Daubechies, De Mol, Fornasier, Tomba, E., Lenti, ...]

To reduce over-smoothness, these methods have been implemented in the context of L^p Lebesgue spaces with $1 < p \leq 2$.

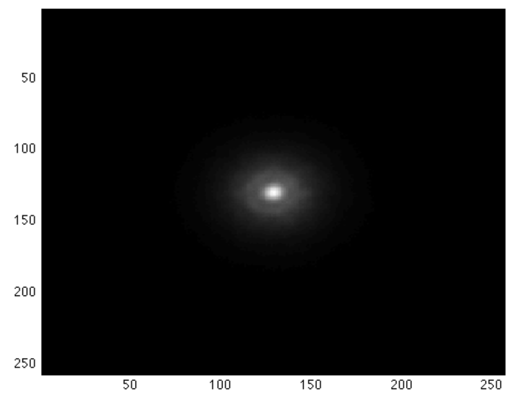
$p >\approx 1$	Low regularization	Good recovery of edges and discontinuities in imaging. Improve the sparsity.
$p \approx 2$	High regularization	Higher stability Over-smoothness

A numerical evidence in L^p Lebesgue space, $1 < p \leq 2$

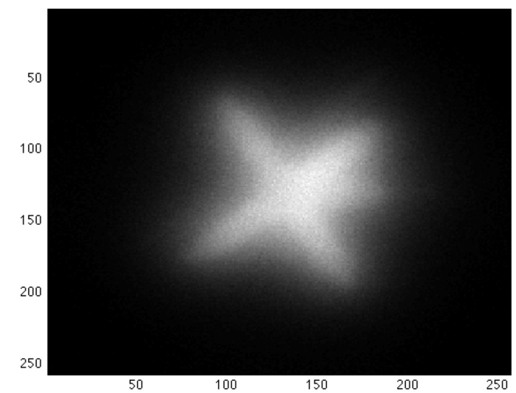
Landweber method (200 iterations)



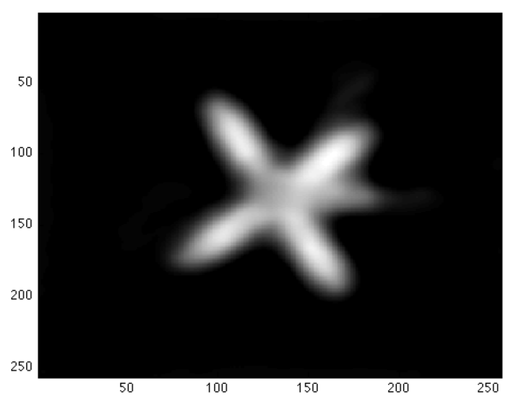
True image x



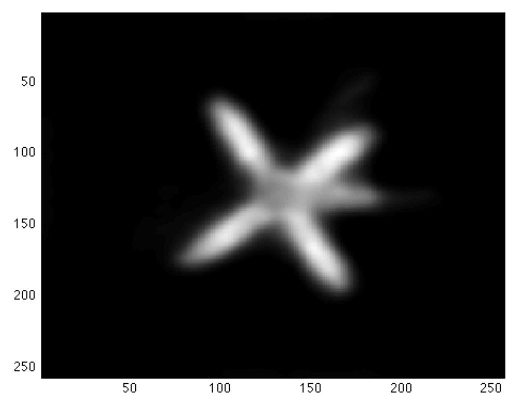
PSF



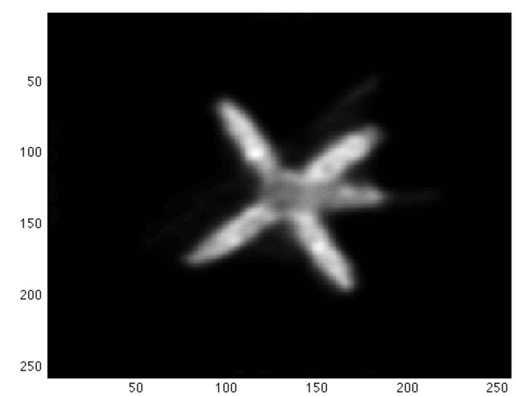
Blurred and noisy image



Hilbert Restoration ($p = 2$)



Banach Restoration ($p = 1.5$)



Banach Restoration ($p = 1.2$)

A “new” framework: variable exponent Lebesgue spaces $L^{p(\cdot)}$

In image restoration, often different regions of the image require different “amount of regularization”.
Setting different levels of 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).

The idea: the ill-posed functional equation $Af = g$ is solved in $L^{p(\cdot)}$ Banach spaces, namely, the variable exponent Lebesgue spaces, a special case of the so-called Musielak-Orlicz functional spaces (first proposed in two seminal papers in 1931 and 1959, but intensively studied just in the last 10 years).

In a variable exponent Lebesgue space, 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 $1 \leq p(\cdot) \leq +\infty$.

This way, different regularization levels on different regions of the image to restore can be automatically and adaptively assigned.

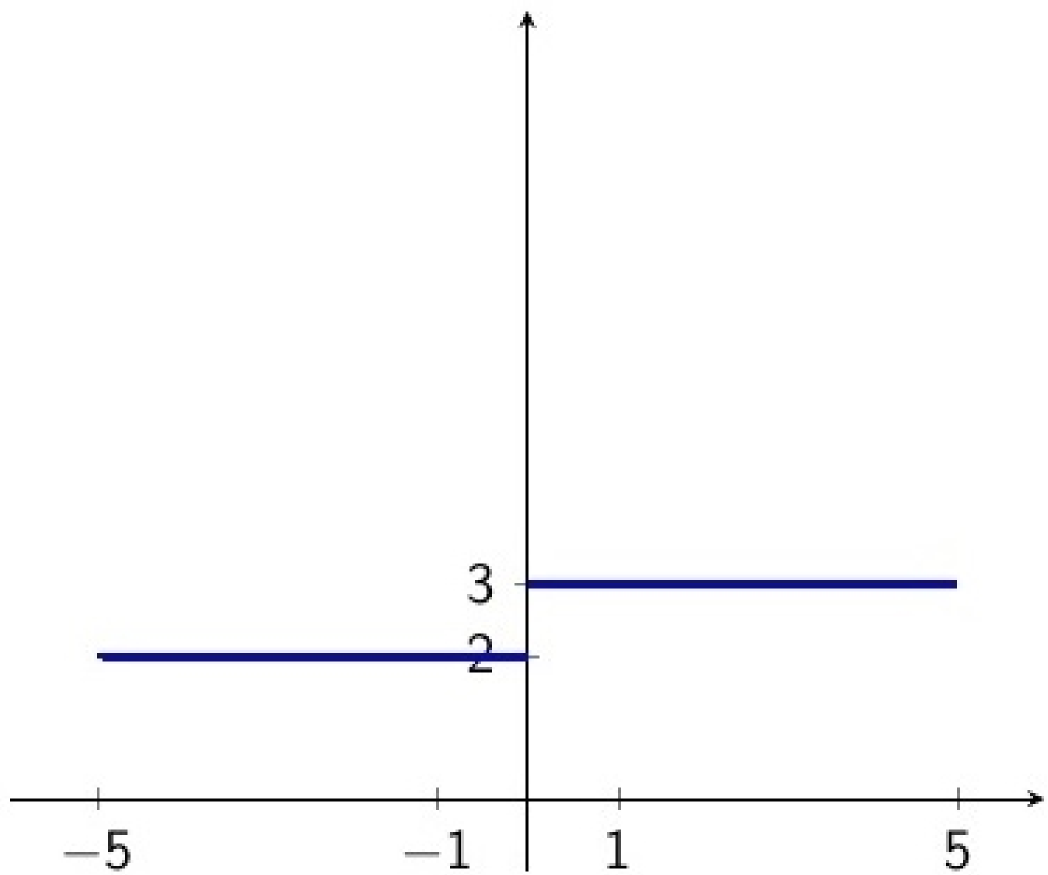
A sketch on variable exponent Lebesgue spaces $L^{p(\cdot)}$

$L^p(\Omega)$	$L^{p(x)}(\Omega)$
$1 \leq p \leq \infty$ p is constant	$p(x) : \Omega \to [1, \infty]$ $p(x)$ is a measurable function
$\Vert f \Vert_p = \left(\int_\Omega \vert f(x) \vert^p \, dx \right)^{1/p}$ $\Vert f \Vert_\infty = \operatorname{ess\,sup} \vert f(x) \vert$	$\Vert f \Vert_{p(\cdot)} = \left(\int_\Omega \vert f(x) \vert^{p(x)} \, dx \right)^{1/???}$ $\ldots$
$f \in L^p(\Omega) \iff \int_\Omega \vert f(x) \vert^p \, dx < \infty$	$f \in L^{p(\cdot)}(\Omega) \iff ???$

In the following, $\Omega_\infty = \{x \in \Omega : p(x) = \infty\}$ has zero measure.

$$\Omega = [-5, 5]$$

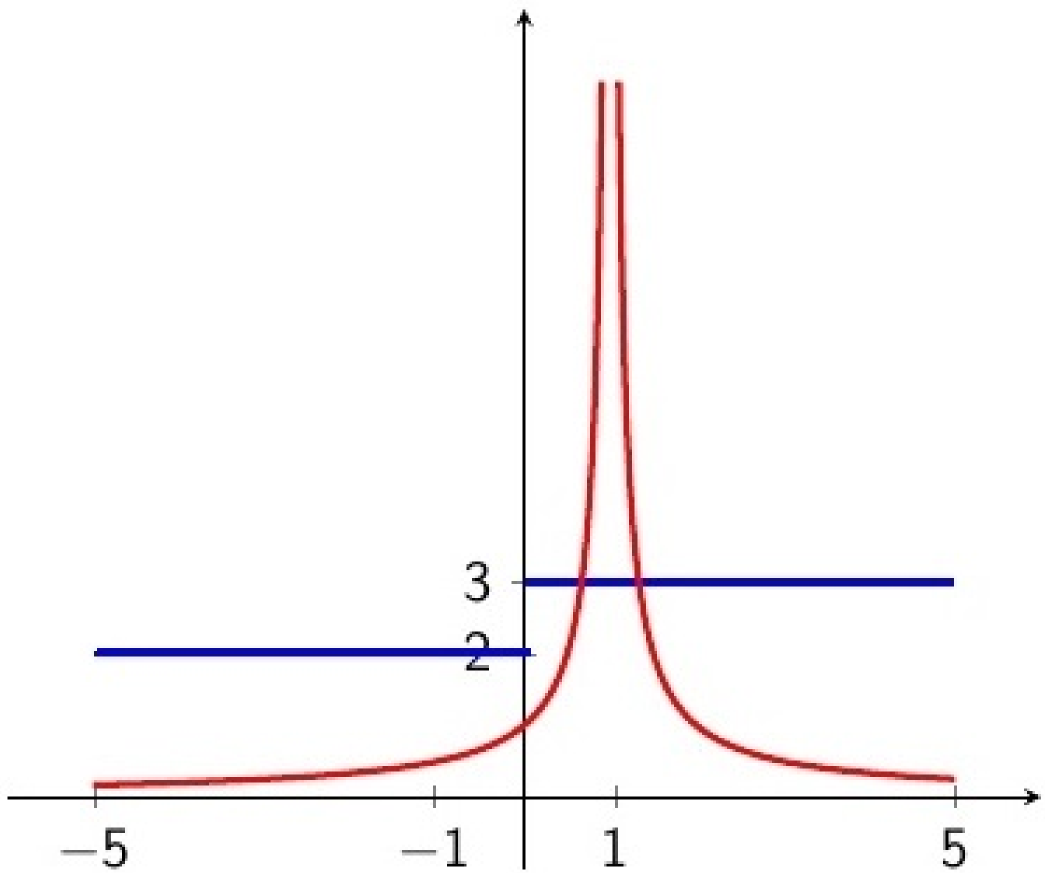
$$p(x) = \begin{cases} 2 & \text{se} \quad -5 \leq x \leq 0 \\ 3 & \text{se} \quad 0 < x \leq 5 \end{cases}$$



$$\Omega = [-5, 5]$$

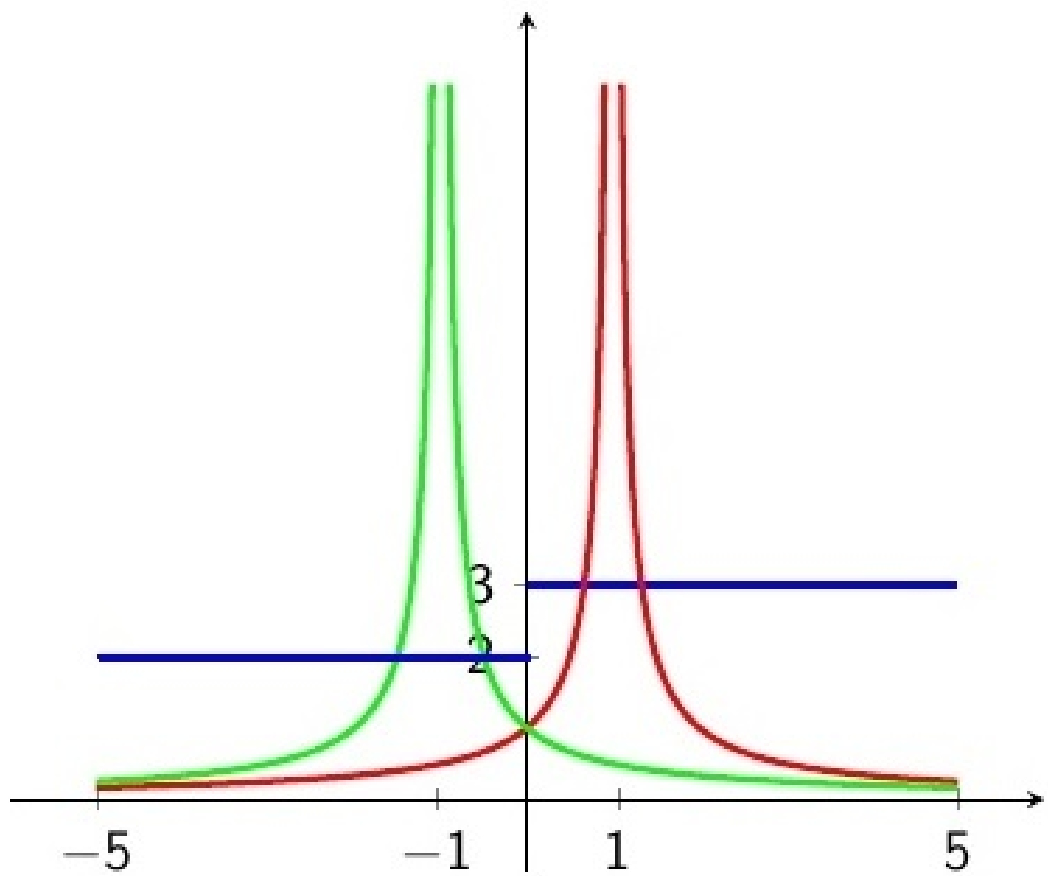
$$p(x) = \begin{cases} 2 & \text{se} \quad -5 \leq x \leq 0 \\ 3 & \text{se} \quad 0 < x \leq 5 \end{cases}$$

$$f(x) = \frac{1}{|x-1|^{1/3}} \notin L^{p(\cdot)}([-5, 5])$$



$$\Omega = [-5, 5] \qquad\qquad p(x) = \begin{cases} 2 & \text{se} \quad -5 \leq x \leq 0 \\ 3 & \text{se} \quad 0 < x \leq 5 \end{cases}$$

$$f(x) = \frac{1}{|x+1|^{1/3}} \in L^{p(\cdot)}([-5,5]) \qquad\qquad f(x) = \frac{1}{|x-1|^{1/3}} \notin L^{p(\cdot)}([-5,5])$$



The norm of variable exponent Lebesgue spaces

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

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.

$$\|f\|_{L^{p(\cdot)}} = \Big(\int_\Omega |f(x)|^{p(x)} dx \Big)^{1/???}.$$

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

$$\rho(f) = \int_\Omega |f(x)|^{p(x)} dx \,,$$

and then obtain the (so called Luxemburg [1955]) norm by solving a 1D minimization problem

$$\|f\|_{L^{p(\cdot)}} = \inf \Big\{ \lambda > 0 : \rho\Big(\frac{f}{\lambda}\Big) \leq 1 \Big\}.$$

The elements of a variable exponent Lebesgue space

$$\rho(f) = \int_\Omega |f(x)|^{p(x)} dx \,,$$

$$\|f\|_{L^{p(\cdot)}} = \inf\left\{\lambda > 0 : \rho\Big(\frac{f}{\lambda}\Big) \le 1\right\}.$$

The Lebesgue space

$$L^{p(\cdot)}(\Omega) = \left\{f : \Omega \to \mathbb{R} \;\mid\; \|f\|_{L^{p(\cdot)}} < \infty\right\}$$

is a Banach space.

In the case of a constant function exponent $p(x) = p$, this norm is exactly the classical one $\|f\|_p$, indeed

$$\rho\Big(\frac{f}{\lambda}\Big) = \int_\Omega \Big(\frac{|f(x)|}{\lambda}\Big)^p dx = \frac{1}{\lambda^p}\int_\Omega |f(x)|^p dx = \frac{1}{\lambda^p}\|f\|_p^p$$

and

$$\inf\left\{\lambda > 0 : \frac{1}{\lambda^p}\|f\|_p^p \le 1\right\} = \|f\|_p$$

Modulus VS Norm

In (classical) L^p, norm and modulus are "the same" apart from a p-root:

$$\|f\|_p < \infty \qquad \Longleftrightarrow \qquad \int_\Omega |f(x)|^p \, dx < \infty$$

In $L^{p(\cdot)}$, norm and modulus are really different:

$$\|f\|_{p(\cdot)} < \infty \qquad \Leftarrow \not\Rightarrow \qquad \rho(f) < \infty$$

Indeed, the following holds

$$\|f\|_{p(\cdot)} < \infty \iff \text{there exist a } \lambda > 0 \text{ s.t. } \rho\Big(\frac{f}{\lambda}\Big) < \infty$$

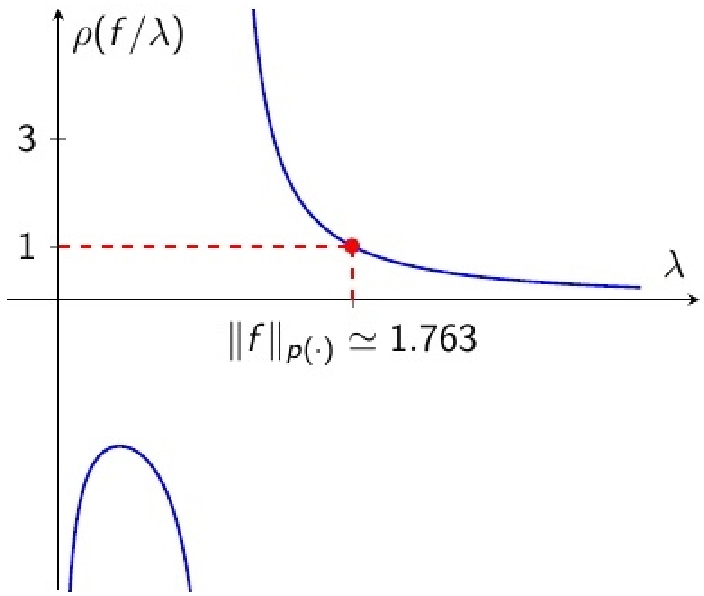
(and notice that λ can be chosen large enough ...).

A simple example of a strange behavior

$$\Omega = [1, \infty) \qquad\qquad f(x) \equiv 1 \qquad\qquad p(x) = x$$

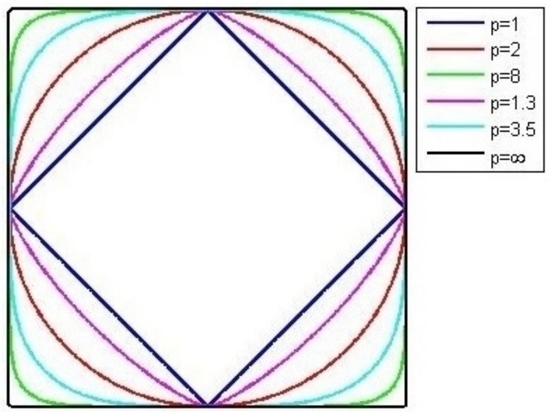
$$\rho(f) = \int_1^\infty 1^x \, dx = \infty \qquad \text{BUT} \qquad \|f\|_{p(\cdot)} \simeq 1.763$$

$$\text{Indeed} \quad \rho(f/\lambda) = \int_1^\infty \left(\frac{1}{\lambda}\right)^x dx = \frac{1}{\lambda \log \lambda} < \infty, \quad \text{for any} \quad \lambda > 1$$

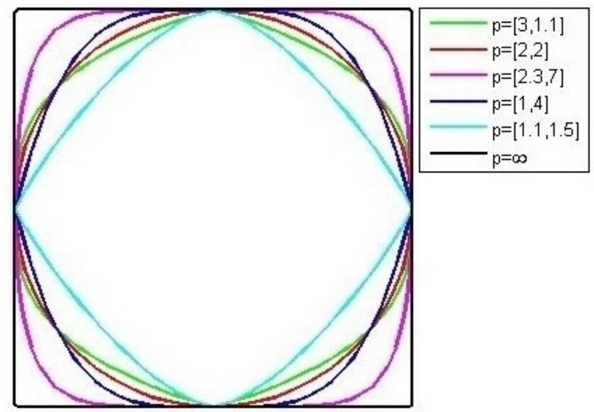


The vector case: the Lebesgue spaces of sequences $l^{p(\cdot)} = (l_n^{p_n})$

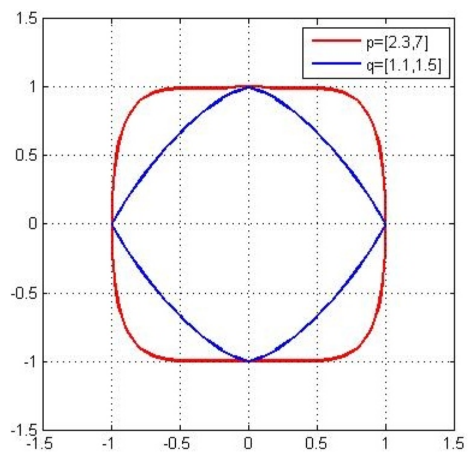
The unit circle of $x = (x_1; x_2)$ in $\mathbb{R}^2$ with variable exponents $p = (p_1; p_2)$.



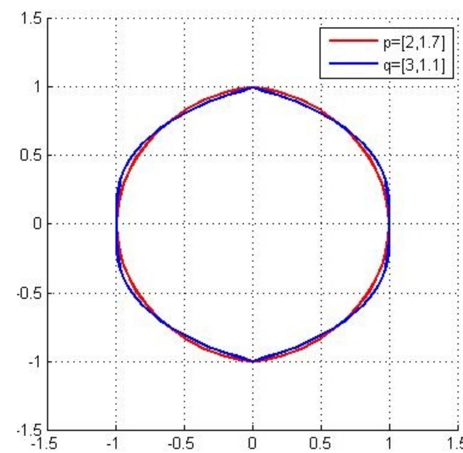
$\|x\|_p$



$\|x\|_{p(\cdot)}$



Inclusion if $(p_1, p_2) \geq (q_1, q_2)$ (as classical)



No inclusion in general

Properties of variable exponent Lebesgue spaces $L^{p(\cdot)}$

$$\text{Let} \quad p_- = \operatorname*{ess\,inf}_{\Omega} |p(x)| \ , \quad \text{and} \quad p_+ = \operatorname*{ess\,sup}_{\Omega} |p(x)| \ .$$

If $p_+ = \infty$, then $L^{p(\cdot)}(\Omega)$ is a "bad" (although very interesting) Banach space, with poor geometric properties (i.e., not useful for our regularization schemes).

If $1 < p_- \leq p_+ < \infty$, then $L^{p(\cdot)}(\Omega)$ is a "good" Banach space, since many properties of classical Lebesgue spaces L^p still hold.

This is the natural framework for our iterative methods in Banach spaces, because:

- $L^{p(\cdot)}$ is uniformly smooth, uniformly convex, and reflexive,
- its dual space is well defined, $\left(L^{p(\cdot)}\right)^* \simeq L^{q(\cdot)}$, where $\frac{1}{p(x)} + \frac{1}{q(x)} = 1$,
- we can define its duality map.

The duality map of the variable exponent Lebesgue space

By extending the duality maps, we can define into $L^{p(\cdot)}$ all the iterative methods developed in L^p (Landweber, Steepest descent, CG, Mann iter.).

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 $p \approx 2$ (even $p > 2$), allows a better restoration of the regions of pixels with the highest intensities.

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

$$p(\cdot) = 1 + [Ag(\cdot) - \min(Ag)]/[\max(Ag) - \min(Ag)]$$

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

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

The conjugate gradient method in $L^{p(\cdot)}$ for image restoration

Let $p(\cdot) = 1 + [Ag(\cdot) - \min(Ag)]/[\max(Ag) - \min(Ag)]$

$$\rho_0^* = -A^* J_r^{L^{p(\cdot)}}(Af_0 - g)$$

For $k = 1, 2, 3, \ldots$

$$\alpha_k = \arg\min_\alpha \tfrac{1}{r} \|A(f_k + \alpha\rho_k) - g\|^r_{L^{p(\cdot)}}$$

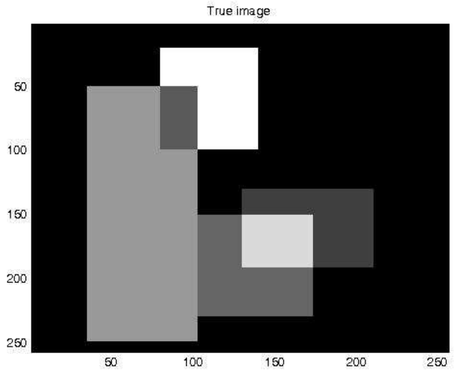
$$f^*_{k+1} = f^*_k + \alpha_k \rho^*_k \qquad f_{k+1} = J_{s'}^{(L^{p(\cdot)})^*}(f^*_{k+1})$$

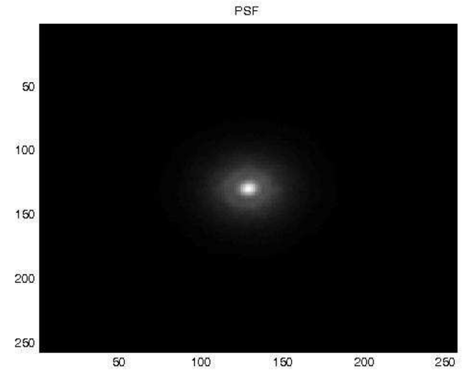
$$\beta_{k+1} = -\gamma \frac{\|Af_{k+1} - g\|^r_{L^{p(\cdot)}}}{\|Af_k - g\|^r_{L^{p(\cdot)}}} \qquad \text{with} \gamma < 1/2$$

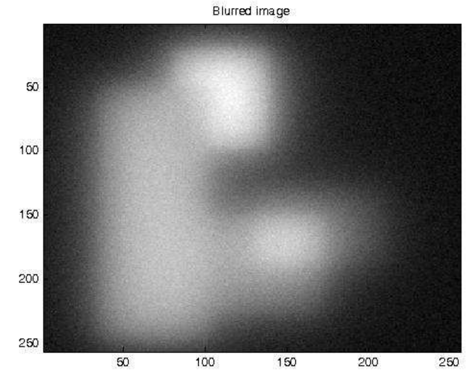
$$\rho^*_{k+1} = -A^* J_r^{L^{p(\cdot)}}(Af_{k+1} - g) + \beta_{k+1}\rho^*_k$$

(and recompute $p(\cdot) = 1 + [f_k(\cdot) - \min(f_k)]/[\max(f_k) - \min(f_k)]$ each m iterations by using the last iteration f_k).

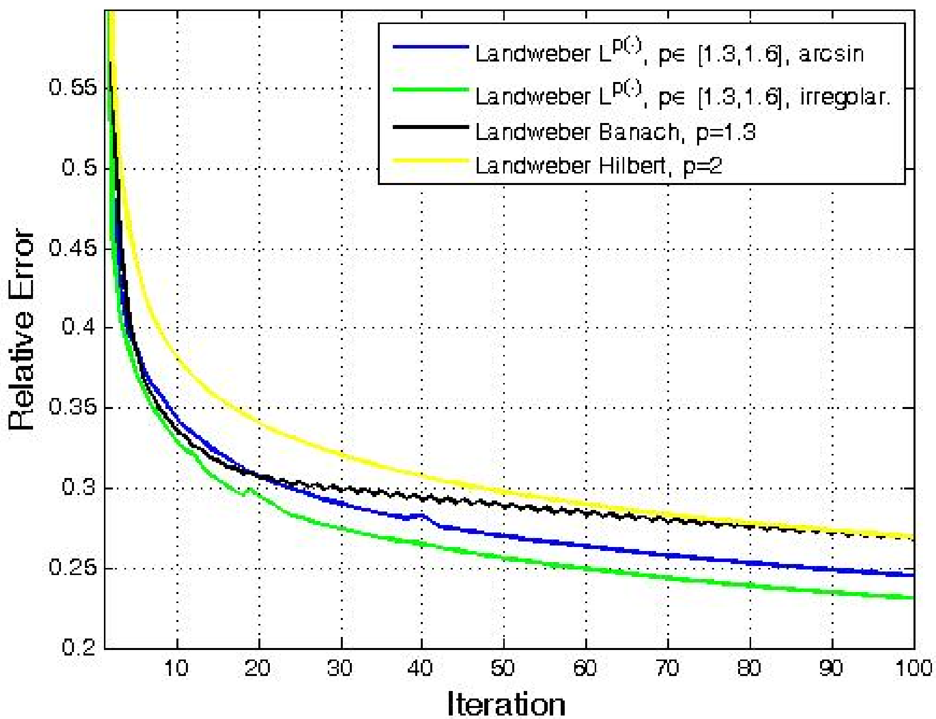
Numerical results

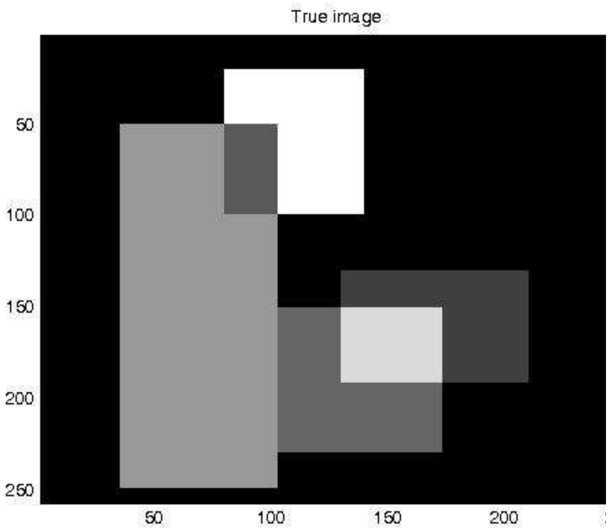




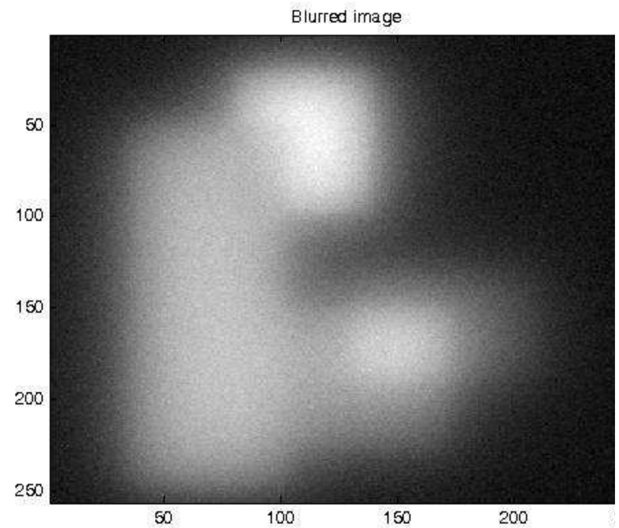


True image | Point Spread Function | Blurred image (noise = 4.7%)

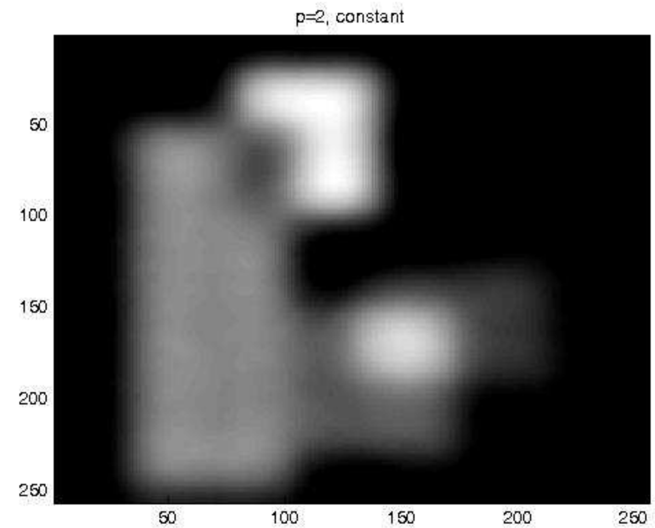




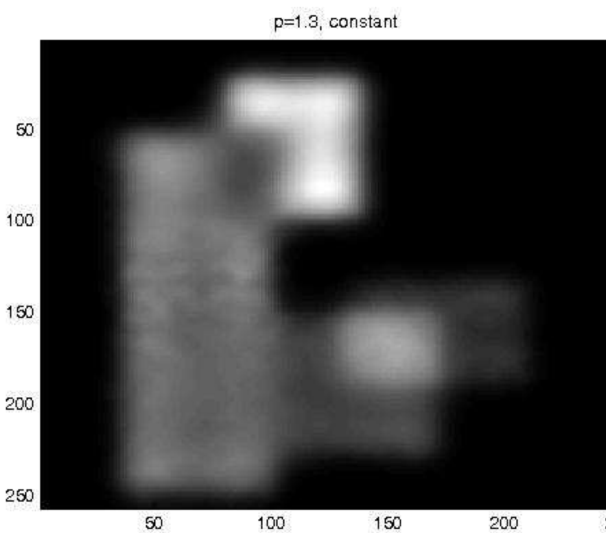
True image



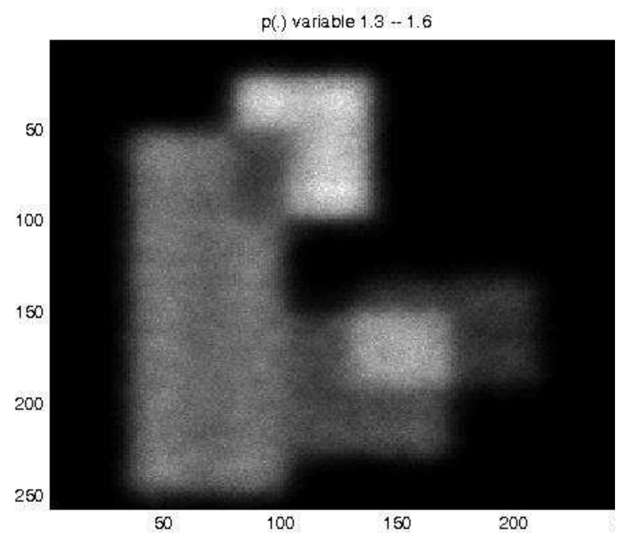
Blurred (noise = 4.7%)



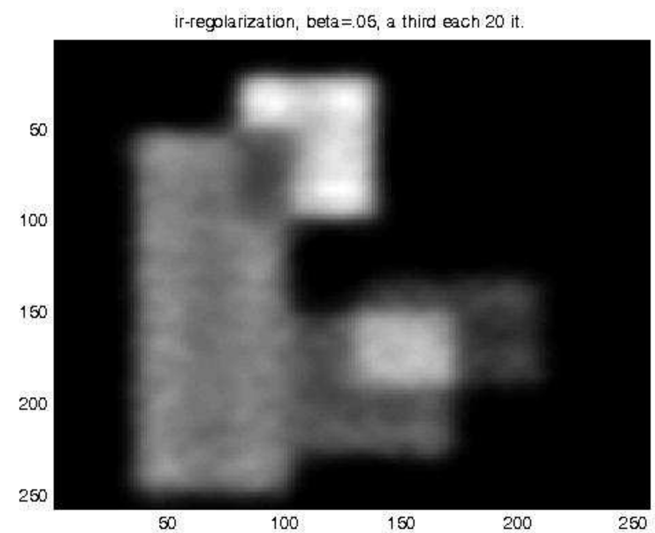
$p = 2 \quad (0.2692)$



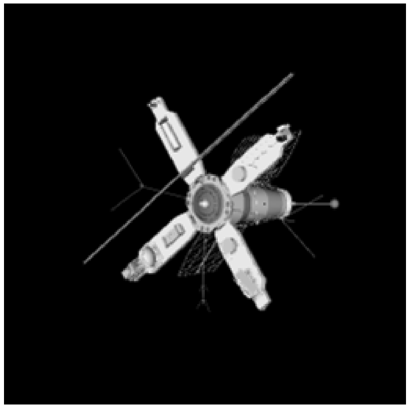
$p = 1.3 \quad (0.2681)$



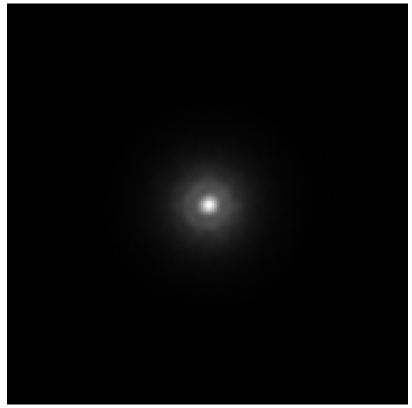
$p = 1.3 - 1.6 \quad (0.2473)$



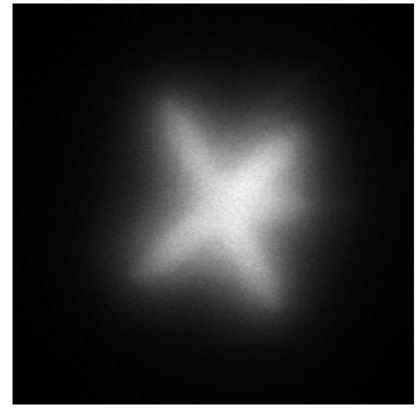
$1.3 - 1.6$ and irreg. (0.2307)



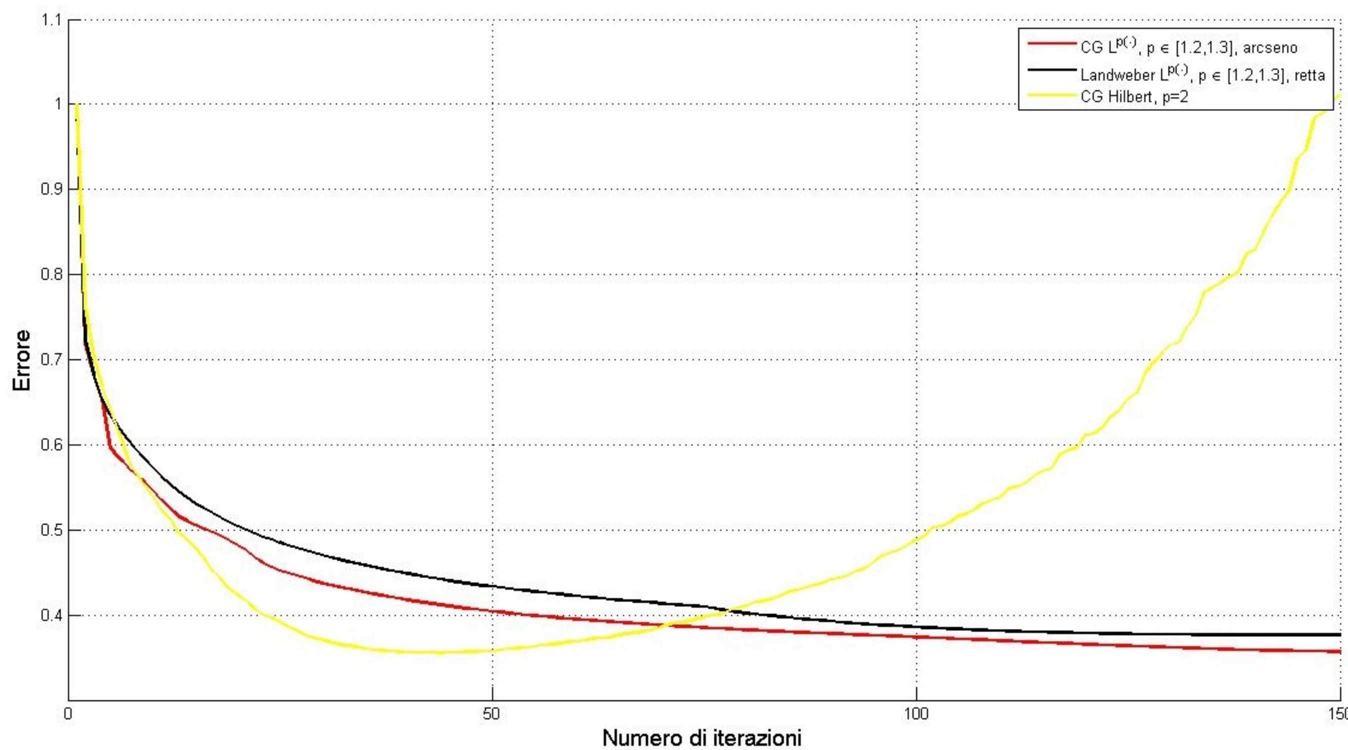
True image

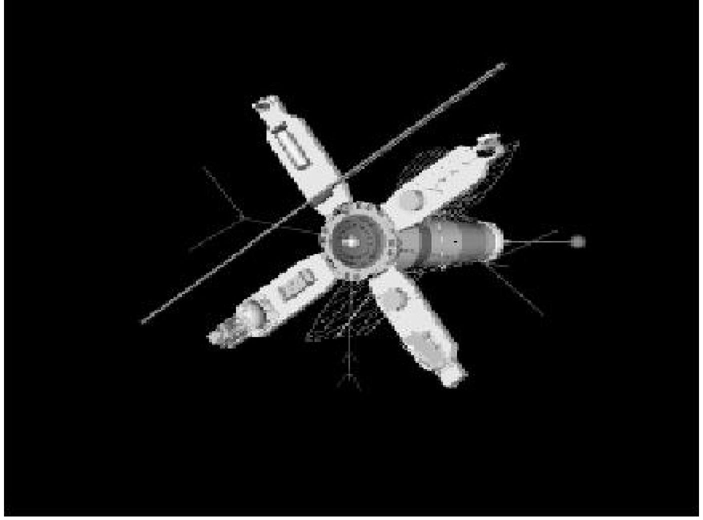


Point Spread Function



Blurred image (noise = 4.7%)

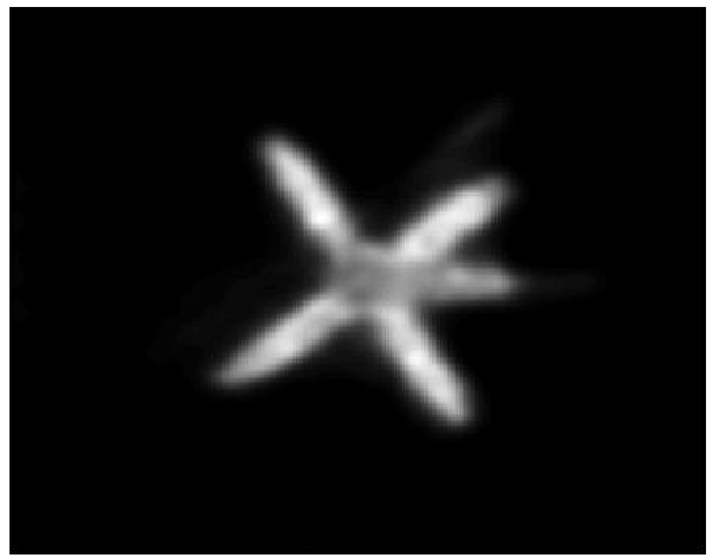




True image

CG $p(\cdot)$ (it. 150; rre: 0.3569)

CG $p = 2$ (it. 45; rre: 0.3557)



Landw. $p(\cdot)$ (it. 150; rre: 0.3766)

Conclusions

- Iterative methods in variable Lebesgue spaces seems to be promising for implementing adaptive “regularization”.

- The framework is new (previously only used in the penalty term as ”weighted total variation” $\int |\nabla(f)|^{p(\cdot} dx$, as proposed in [1997 by T. Chan et al]).

- Theoretical convergence is under analysis.

Thank you for your attention.

References

[1] M. O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, F. Lenzen,
Variational Methods in Imaging, Springer, Berlin, 2008.

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

[3] F. Schöpfer, A.K. Louis, T. Schuster,
Nonlinear iterative methods for linear ill-posed problems in Banach spaces,
Inverse Problems, vol. 22, pp. 311–329, 2006.

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

[5] P. Matei,
A nonlinear eigenvalue problem for the gener. Laplacian on Sobolev spaces with variable exponent,
Romanian J. of Mathematics and Comp. Science, vol. 2, pp. 73–82, 2012.

[6] C. Estatico, S. Gratton, F. Lenti, D. Titley-Peloquin,
A conjugate gradient based method for p-norm minimization,
submitted, 2016.