Multilevel regularization for image deblurring problems

Marco Donatelli
joint with S. Serra-Capizzano

University of Insubria
Department of Physics and Mathematics

ICIAM 2007 Zurich, July 2007

Outline

Image restoration with Boundary Conditions

Using Boundary Conditions (BCs), the restored image $\mathbf{f}$ is obtained solving: (in some way ...)

$$A\mathbf{f} = \mathbf{g} + \boldsymbol{\xi}$$

- $\mathbf{g}$ = blurred image,
- ξ = noise (random vector),
- A = two-level matrix depending on the point spread function (PSF) and the BCs.

Coefficient matrix structure

The matrix-vector product computed in $O(n^2 \log(n))$ ops for $n \times n$ images while the inversion costs $O(n^2 \log(n))$ ops only in the periodic case.

BCs	A
Dirichlet	Toeplitz
periodic	circulant
Neumann (reflective)	Toeplitz + Hankel
anti-reflective	Toeplitz + Hankel

Coefficient matrix structure

The matrix-vector product computed in $O(n^2 \log(n))$ ops for $n \times n$ images while the inversion costs $O(n^2 \log(n))$ ops only in the periodic case.

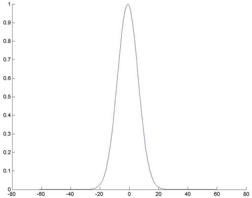
BCs	A
Dirichlet	Toeplitz
periodic	circulant
Neumann (reflective)	Toeplitz + Hankel
anti-reflective	Toeplitz + Hankel

If the PSF is symmetric with respect to each direction:

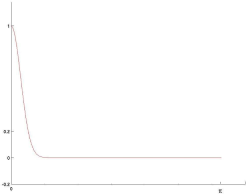
BCs	A
Neumann (reflective)	DCT III
anti-reflective	DST I + low-rank

Generating function of PSF

- The eigenvalues of $A(z)$ are about a uniform sampling of z.



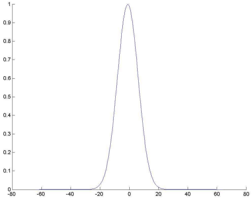
PSF



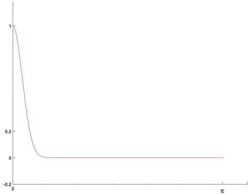
Generating function $z(x)$

Generating function of PSF

- The eigenvalues of $A(z)$ are about a uniform sampling of z.



PSF



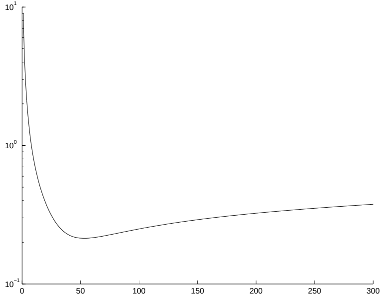
Generating function $z(x)$

- The ill-conditioned subspace is mainly constituted by the middle/high frequencies.

Semi-convergence of iterative regularization methods

Some iterative methods (CGNE, . . .) have regularization properties:

- They firstly reduce the algebraic error in the low frequencies (well-conditioned subspace).
- When they arrive to reduce the algebraic error in the high frequencies then the restoration error increases because of the noise.



Multigrid structure

Multigrid Idea

Project the system in a subspace, solve the resulting system in this subspace and interpolate the solution in order to improve the previous approximation.

Multigrid structure

Multigrid Idea

Project the system in a subspace, solve the resulting system in this subspace and interpolate the solution in order to improve the previous approximation.

Multigrid components

The Multigrid combines two iterative methods:

Smoother: a classic iterative method,

Multigrid structure

Multigrid Idea

Project the system in a subspace, solve the resulting system in this subspace and interpolate the solution in order to improve the previous approximation.

Multigrid components

The Multigrid combines two iterative methods:

Smoother: a classic iterative method,

Coarse Grid Correction: projection, solution of the restricted problem, interpolation.

Multigrid structure

Multigrid Idea

Project the system in a subspace, solve the resulting system in this subspace and interpolate the solution in order to improve the previous approximation.

Multigrid components

The Multigrid combines two iterative methods:

- **Smoother:** a classic iterative method,
- **Coarse Grid Correction:** projection, solution of the restricted problem, interpolation.

At the lower level(s) it works on the error equation!

The algorithm

Two-Grid Methods (TGM)

The j-th iteration for the system $A\mathbf{x} = \mathbf{b}$:

$$
\begin{aligned}
&\textbf{(1)}\quad \tilde{\mathbf{x}} = \mathit{Smooth}(A, \mathbf{x}^{(j)}, \mathbf{b}, \nu) \\
&\textbf{(2)}\quad \mathbf{r}_1 = \mathbf{P}(\mathbf{b} - A\tilde{\mathbf{x}}) \\
&\textbf{(3)}\quad A_1 = \mathbf{P}A\mathbf{P}^T \\
&\textbf{(4)}\quad \mathbf{e}_1 = A_1^{-1}\mathbf{r}_1 \\
&\textbf{(5)}\quad \mathbf{x}^{(j+1)} = \mathbf{x}^{(j)} + \mathbf{P}^T\mathbf{e}_1
\end{aligned}
$$

The algorithm

Two-Grid Methods (TGM)

The j-th iteration for the system $A\mathbf{x} = \mathbf{b}$:

$$
\begin{aligned}
&(1)\quad \tilde{\mathbf{x}} = Smooth(A, \mathbf{x}^{(j)}, \mathbf{b}, \nu)\\
&(2)\quad \mathbf{r}_1 = \mathbf{P}(\mathbf{b} - A\tilde{\mathbf{x}})\\
&(3)\quad A_1 = \mathbf{P}A\mathbf{P}^T\\
&(4)\quad \mathbf{e}_1 = A_1^{-1}\mathbf{r}_1\\
&(5)\quad \mathbf{x}^{(j+1)} = \mathbf{x}^{(j)} + \mathbf{P}^T\mathbf{e}_1
\end{aligned}
$$

Multigrid (MGM): the step **(4)** becomes a recursive application of the algorithm.

Algebraic Multigrid

1. The AMG uses information on the coefficient matrix and no geometric information on the problem.

Algebraic Multigrid

1. The AMG uses information on the coefficient matrix and no geometric information on the problem.
2. We use the Galerkin approach
   - $P_i = R_i^T$
   - $A_{i+1} = RA_iR_i^T$

   where R_i and P_i are the restriction and the prolongation operators at the level i, respectively.

Algebraic Multigrid

1. The AMG uses information on the coefficient matrix and no geometric information on the problem.
2. We use the Galerkin approach
   - $P_i = R_i^T$
   - $A_{i+1} = RA_iR_i^T$

   where R_i and P_i are the restriction and the prolongation operators at the level i, respectively.
3. The smoother is a classic iterative regularization method: in the initial iterations it is not able to reduce effectively the error in the ill-conditioned subspace.

Algebraic Multigrid

1. The AMG uses information on the coefficient matrix and no geometric information on the problem.
2. We use the Galerkin approach
   - $P_i = R_i^T$
   - $A_{i+1} = RA_iR_i^T$

   where R_i and P_i are the restriction and the prolongation operators at the level i, respectively.
3. The smoother is a classic iterative regularization method: in the initial iterations it is not able to reduce effectively the error in the ill-conditioned subspace.
4. Coarser grid of size 8×8 independent of the size of the finer grid.

Image deblurring and Multigrid

- In the image deblurring the ill-conditioned subspace is related to high frequencies, while the well-conditioned subspace is generated by low frequencies.

Image deblurring and Multigrid

- In the image deblurring the ill-conditioned subspace is related to high frequencies, while the well-conditioned subspace is generated by low frequencies.
- In order to obtain a fast convergence the algebraic multigrid projects in the high frequencies where the noise "lives" $\Longrightarrow$ noise explosion already at the first iteration (it requires Tikhonov regularization [Donatelli, NLAA, 12 (2005), pp. 715–729]).

Image deblurring and Multigrid

- In the image deblurring the ill-conditioned subspace is related to high frequencies, while the well-conditioned subspace is generated by low frequencies.
- In order to obtain a fast convergence the algebraic multigrid projects in the high frequencies where the noise "lives" $\Longrightarrow$ noise explosion already at the first iteration (it requires Tikhonov regularization [Donatelli, NLAA, 12 (2005), pp. 715–729]).
- In this case the low-pass filter projects in the well-conditioned subspace (low frequencies) $\Longrightarrow$ it is slowly convergent but it can be a good iterative regularizer.

Restriction operator

$R_i = K_{N_i} \mathcal{A}_{N_i}(p)$:

- $K_{N_i} \in \mathbb{R}^{\frac{N_i}{4} \times N_i}$ is the cutting matrix that preserves the structure at the lower level.

circulant	**Toeplitz & DST − I**	**DCT − III**
$\begin{bmatrix} 1 & 0 & & & & \\ & & 1 & 0 & & \\ & & & \ddots & \ddots & \\ & & & & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & & & \\ & & 0 & 1 & 0 & \\ & & & \ddots & \ddots & \ddots \\ & & & & 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 0 & & & \\ & & 1 & 1 & 0 & \\ & & & \ddots & \ddots & \ddots \\ & & & & 0 & 1 & 1 \end{bmatrix}$

Restriction operator

$R_i = K_{N_i}\mathcal{A}_{N_i}(p)$:

- $K_{N_i} \in \mathbb{R}^{\frac{N_i}{4} \times N_i}$ is the cutting matrix that preserves the structure at the lower level.

circulant	**Toeplitz & DST − I**	**DCT − III**
$\begin{bmatrix} 1\ 0 & & & \\ & 1\ 0 & & \\ & & \ddots\ \ddots & \\ & & & 1\ 0 \end{bmatrix}$	$\begin{bmatrix} 0\ 1\ 0 & & & \\ & 0\ 1\ 0 & & \\ & & \ddots\ \ddots\ \ddots & \\ & & & 0\ 1\ 0 \end{bmatrix}$	$\begin{bmatrix} 1\ 1\ 0 & & & \\ & 1\ 1\ 0 & & \\ & & \ddots\ \ddots\ \ddots & \\ & & & 0\ 1\ 1 \end{bmatrix}$

- $p(x, y)$ is the generating function of the projector, which selects the subspace where to project the linear system.

$$\text{2D} \quad \leftrightarrow \quad p(x, y) = (1 + \cos(x))(1 + \cos(y))$$

If we have an *iterative regularization method* we can improve its regularizing properties and/or accelerate its convergence using it as *smoother* in a Multigrid algorithm.

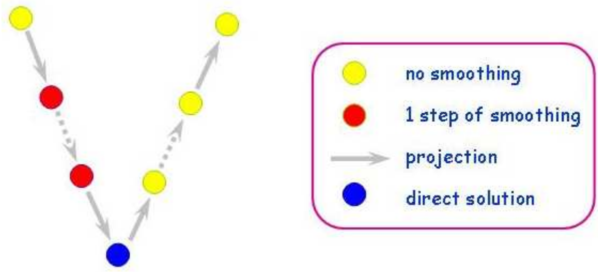
If we have an *iterative regularization method* we can improve its regularizing properties and/or accelerate its convergence using it as *smoother* in a Multigrid algorithm.

Two Level algorithm

1. No smoothing at the finer level
2. At the coarser level to apply one step of the smoother instead of to solve directly the linear system

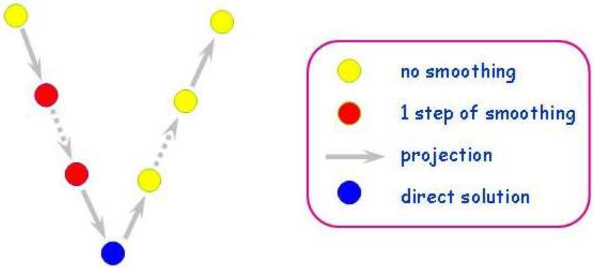
Multigrid regularization (applying recursively the TL)

V-cycle



Multigrid regularization (applying recursively the TL)

V-cycle



Using a larger number of recursive calls (e.g. W-cycle), the algorithm "works" more in the well-conditioned subspace, but it is more difficult to define an early stopping criterium.

Computational Cost

Assumptions: $n \times n$ images and $m \times m$ PSFs with $m \ll n$.

- Let $S(n)$ be the computational cost of one smoother iteration.
- The computational cost of one iteration of our multigrid regularization method with γ recursive calls is

$$C(\gamma, n) \approx \begin{cases} \frac{1}{3}S(n), & \gamma = 1 \\ S(n), & \gamma = 2 \\ 3S(n), & \gamma = 3 \end{cases}$$

- if $m \approx n$ then $S(n) = O(n^2 \log(n))$.

Remarks on the TL regularization

Theoretical analysis

- Intuitively: the regularization properties of the smoother are preserved since it is combined with a low-pass filter.

Remarks on the TL regularization

Theoretical analysis

- Intuitively: the regularization properties of the smoother are preserved since it is combined with a low-pass filter.
- From the filter factor analysis of the Two Level (TL) method we have:
  - a condition to choose p s.t. it is a regularizer
  - in such case it improves the regularization properties of the iterative method used as smoother.

Remarks on the TL regularization

Theoretical analysis

- Intuitively: the regularization properties of the smoother are preserved since it is combined with a low-pass filter.
- From the filter factor analysis of the Two Level (TL) method we have:
    - a condition to choose p s.t. it is a regularizer
    - in such case it improves the regularization properties of the iterative method used as smoother.
- In the noise free case the TL method does not compute the exact solution.

An airplane

- Periodic BCs
- Gaussian PSF (A spd)
- noise = 1%

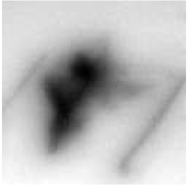
Original Image

An airplane

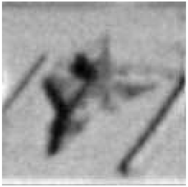
- Periodic BCs
- Gaussian PSF (A spd)
- noise = 1%

Original Image

Inner part 128×128



Observed image



Restored with MGM

Restoration error: noise = 1%

$e_j = \|\mathbf{f} - \mathbf{f}^{(j)}\|_2 / \|\mathbf{f}\|_2$ restoration error at the j-*th* iteration.

Restoration error: noise $= 1\%$

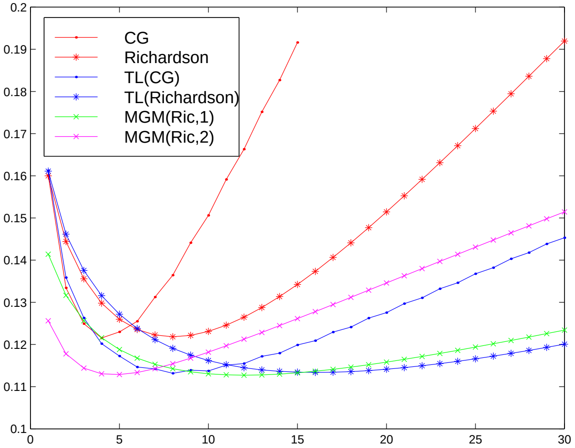
$e_j = \|\mathbf{f} - \mathbf{f}^{(j)}\|_2 / \|\mathbf{f}\|_2$ restoration error at the j-*th* iteration.

Minimum restoration error

Method	$\min_{j=1,\ldots} (e_j)$	$\arg\min_{j=1,\ldots} (e_j)$
CG	0.1215	4
Richardson	0.1218	8
TL(CG)	0.1132	8
TL(Rich)	0.1134	16
MGM(Rich, 1)	0.1127	12
MGM(Rich, 2)	0.1129	5
CGNE	0.1135	178
RichNE	0.1135	352

Restoration error: noise $= 1\%$

Relative error vs. number of iterations



Noise = 10%

For CG and Richardson it is better to resort to normal equations.

Noise $= 10\%$

For CG and Richardson it is better to resort to normal equations.

Minimum restoration error

Method	$\min_{j=1,\dots} (e_j)$	$\arg\min_{j=1,\dots} (e_j)$
CGNE	0.1625	30
RichNE	0.1630	59
TL(CGNE)	0.1611	48
TL(RichNE)	0.1613	97
MGM(RichNE,1)	0.1618	69
MGM(RichNE,2)	0.1621	26
MGM(Rich,1)	0.1648	3
MGM(Rich,2)	0.1630	1

Possible generalizations

- Include nonnegativity constraints.
- Improve the projector:

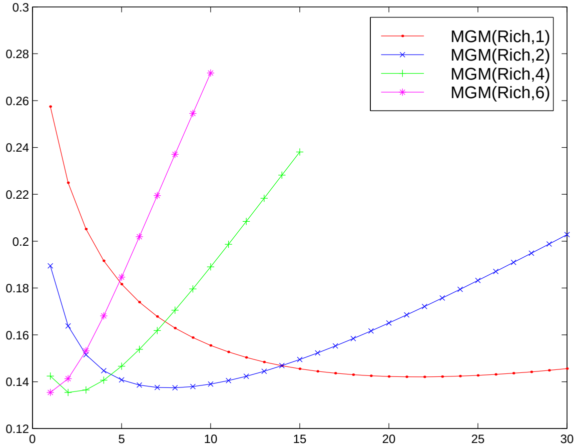
$$p(x, y) = (1 + \cos(x))^\alpha (1 + \cos(y))^\alpha, \qquad \alpha \in \mathbb{N}^+.$$

- The γ regularization:

  varying γ, the proposed multigrid is a direct (one step) regularization method with regularization parameter γ.

  The computational cost increases with γ but not so much (e.g. $\gamma = 8 \Rightarrow O(N^{1.5})$ where $N = n^2$).

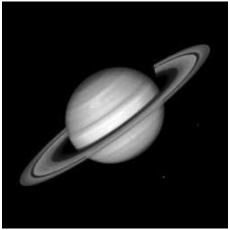
The γ regularization (the airplane with noise = 1%)



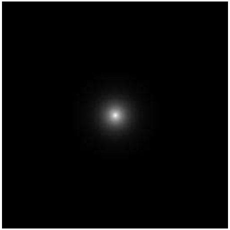
Varying γ, the proposed multigrid is a direct (one step) regularization method with regularization parameter γ.

An astronomic example with nonnegativity constraint

- Periodic BCs and dark background (exact)
- Gaussian PSF + noise = 5%



Original Image



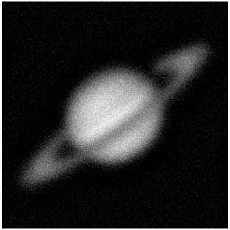
PSF

Observed image

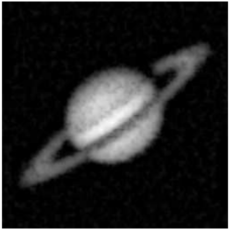
Minimum restoration error

Method	$\min_{j=1,\ldots} (e_j)$	$\arg\min_{j=1,\ldots} (e_j)$
CG^+	0.2268	4
$Rich^+$	0.2298	9
$MGM(Rich^+,1)$	0.1736	17
$MGM(Rich,1)^+$	0.1600	15
$MGM(Rich^+,1)^+$	0.1556	27
$MGM(Rich^+,2)^+$	0.1530	12
$RichNE^+$	0.1419	2735
$CGNE^+$	0.1419	885
$MGM(CGNE^+,2)^+$	0.1389	109
$MGM(CGNE^+,3)^+$	0.1388	45

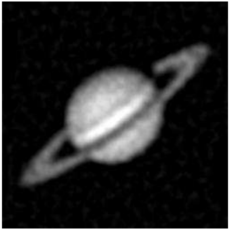
Restored images



Rich^+ (9 iter.)

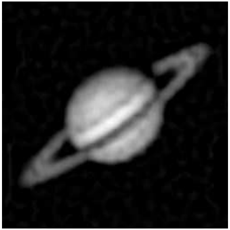


$\text{MGM}(\text{Rich}^+,2)^+$ (12 iter.)

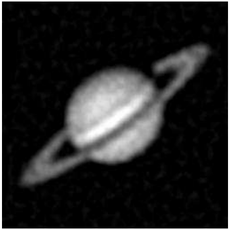


CGNE^+ (885 iter.)

Restored images



MGM(CGNE$^+$,2)$^+$ (109 iter.)



CGNE$^+$ (885 iter.)

Conclusions

- It is a general framework which can be used to improve the regularization properties of an iterative regularizing method.

Conclusions

- It is a general framework which can be used to improve the regularization properties of an iterative regularizing method.
- It is fast and usually it obtains a good restored image also without resorting to normal equations.

Conclusions

- It is a general framework which can be used to improve the regularization properties of an iterative regularizing method.
- It is fast and usually it obtains a good restored image also without resorting to normal equations.

Future work:

Combination with techniques for edge preserving.

Conclusions

- It is a general framework which can be used to improve the regularization properties of an iterative regularizing method.
- It is fast and usually it obtains a good restored image also without resorting to normal equations.

Future work:

Combination with techniques for edge preserving.

Reference

M. Donatelli and S. Serra Capizzano, *On the regularizing power of multigrid-type algorithms*, SIAM J. Sci. Comput., 27–6 (2006) pp. 2053–2076.