Filter factor analysis of an iterative multilevel regularizing method

Marco Donatelli

Department of Physics and Mathematics
University of Insubria

Outline

Outline

Image restoration with Boundary Conditions

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

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

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

The PSF is the observation of a single point (e.g., a star in astronomy).

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.

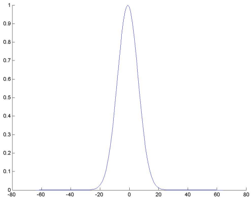
BCs	A
Dirichlet	Toeplitz
periodic	circulant
Neumann (reflective)	Toeplitz + Hankel
anti-reflective	Toeplitz + Hankel

If the PSF is symmetric with respect to each direction:

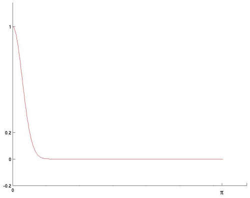
BCs	A
Neumann (reflective)	DCT III
anti-reflective	DST I + low-rank

Generating function of PSF

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



PSF



Generating function $z(x)$

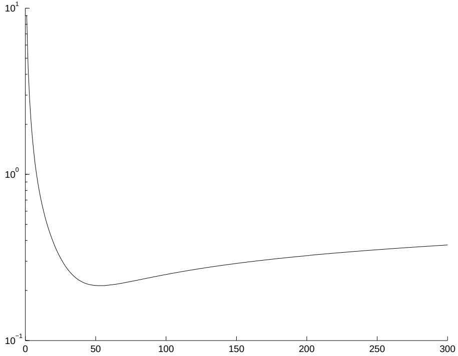
- The ill-conditioned subspace is mainly constituted by the middle/high frequencies.

Iterative regularization methods

Semi-convergence behavior

Some iterative methods (Landweber, CGNE, . . .) have regularization properties: the restoration error firstly decreases and then increases.

Example



Outline

The algorithm

The choices

1. We apply only the pre-smoother simply called smoother.
2. Let R_i and P_i be the restriction and the prolongation operators at the level i, respectively.
3. We use the Galerkin approach
   - $P_i = R_i^T$
   - $A_{i+1} = RA_iR_i^T$
4. Coarser grid of size 8×8 independent of the size of the finer grid.

The Algebraic Multigrid (AMG)

- The AMG uses only information on the coefficient matrix.
- Different classic smoothers have similar behavior:
  in the initial iterations they are not able to reduce effectively the error in the subspace generated by the eigenvectors associated to small eigenvalues (ill-conditioned subspace)

$$\Downarrow$$

- To obtain a fast solver, the restriction is chosen in order to project the error equation in such subspace.

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.

Multigrid for structured matrices

Preserve the structure

- In order to apply recursively the MGM, it is necessary to keep the same structure at each level (Toeplitz, ...).
- For every structure arising from the proposed BCs, there exist projectors that preserve the same structure.

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

- $K_{N_i} \in \mathbb{R}^{\frac{N_i}{4} \times N_i}$ is the cutting matrix that preserves the structure at the lower level.
- $p(x, y)$ is the generating function of the projector, which selects the subspace where to project the linear system.

Multigrid, structured matrices, and images

The cutting matrix K_{n_i} in 1D

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

Low-pass filter: Low frequencies projection $\Rightarrow$ noise reduction

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

$\searrow$ Full weighting $\quad\quad$ $\nearrow$ Bilinear interpolation

Iterative multigrid regularization

The Multigrid as an iterative regularization method

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.

Regularization

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

Two-Level (TL) regularization

Idea: project into the low frequencies and then apply an iterative regularization method.

TL as a specialization of TGM

Smoother: iterative regularization method

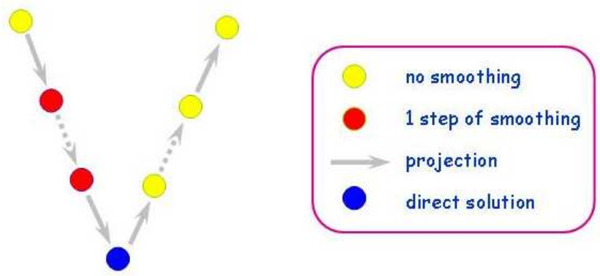
Projector: low-pass filter

TL 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



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 $S(n) = O(n^2 \log(n))$.

Filter factor of the Landweber method

- Imposing P-BCs $A = \mathcal{C}_n(z)$: A is a circulant matrix of size n generated by the function z.
- $A = F_n D_n(z) F_n^H$, where $F_n = [\mathrm{e}^{\mathrm{i} j x_k}]_{k,j=0}^{n-1}/\sqrt{n}$ is the DFT matrix and $D_n(z) = \mathrm{diag}([f(x_k)]_{k=0}^{n-1})$ with $x_k = \frac{2\pi k}{n}$.
- Taking $\mathbf{x}_0 = 0$ the jth approximation of $\mathbf{f}$ is

$$\mathbf{x}_j = F_n \sum_{i=0}^{j-1} (I - D_n(|z|^2))^i D_n(\bar{z}) F_n^H \mathbf{b} \;=\; C_n(\phi_j) C_n^{-1}(z) \mathbf{b}$$

where $\phi_j(x) = 1 - (1 - |z(x)|^2)^j$, $x \in (0, 2\pi]$ is the filter factor.

Filter factor of the TL method

- For TL with Landweber as smoother $\mathbf{x}_j = B_n \mathbf{b}$ with

$$B_n = \mathcal{C}_n(p) K_n^T \mathcal{C}_{\frac{n}{2}}(\hat{g}) K_n \mathcal{C}_n(r),$$

  where $\hat{g}(x) = \frac{1-(1-|\hat{z}(x)|^2)^j}{\hat{z}(x)}$, $x \in (0, 2\pi]$, K_n is the cutting matrix and r, p and $\hat{z}$ are restriction, prolongation and PSF function at the coarser level respectively.

- $B_n = F_n \Pi_n^T W_n \Pi_n F_n^H$, where Π_n is a permutation matrix and W_n is the diagonal block matrix of size $(n/2) \times (n/2)$ with blocks of dimension 2×2. For $k = 0, \ldots, n/2-1$, the k-th diagonal block is given by

$$W_n^{(k)} = \frac{1}{2} \hat{g}(x_{2k}) \left[\begin{array}{c} p(x_k) \\ p(x_{(k+n/2)}) \end{array} \right] \left[\begin{array}{cc} r(x_k) & r(x_{(k+n/2)}) \end{array} \right].$$

Filter factor of the TL method 2

- The block $W_n^{(k)}$ has rank 1 and the nontrivial null eigenvalue λ_k is

$$\lambda_k = \frac{1}{2}\hat{g}(x_{2k})\left((pr)(x_k) + (pr)(x_{(k+n/2)})\right).$$

- The eigenvector associated to the null eigenvalue is

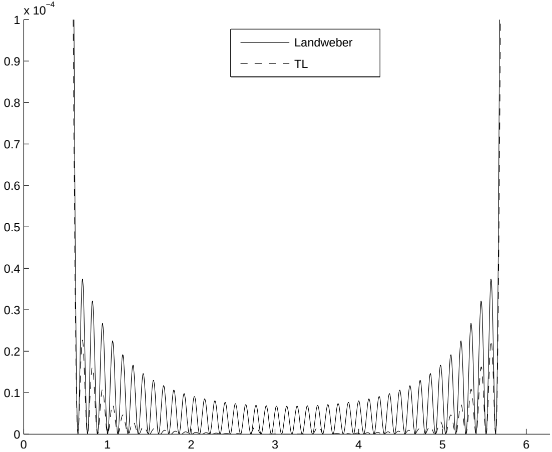
$$\frac{r(x_k)}{r(x_{(k+n/2)})} F_n^{(k+n/2)} - F_n^{(k)}.$$

This should be an high frequency (to filtering) $\Rightarrow$ it provides a condition to choose r: e.g. nonnegative and decreasing in $[0, pi]$.

- The eigenvector associated to λ_k defines an analogous condition for p.

Comparison TL vs Landweber

Focus on the high frequencies for the filter factors of TL and Landweber for $j = 1000$



Noise $\longrightarrow 0$?

In the noise free case the TL method does not compute the exact solution.

How to recover the high frequencies in the noise free case is a work in progress . . .

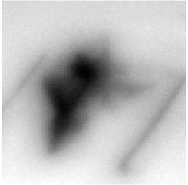
Outline

An airplane

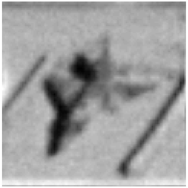
- Periodic BCs
- Gaussian PSF (A spd)
- noise $= 1\%$

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.

Minimum restoration error

Method	$\min_{j=1,\dots} (e_j)$	$\arg\min_{j=1,\dots} (e_j)$
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

Relative error vs. number of iterations

Noise $= 10\%$

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

Minimum restoration error

Method	$\min_{j=1,\ldots} (e_j)$	$\arg\min_{j=1,\ldots} (e_j)$
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

Relative error vs. number of iterations

Outline

Possible generalizations

- Include the 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 $N = n^2$).

Summarizing . . . multigrid regularization method

- It is a general framework which can be used to improve the regularization properties of an iterative regularizing method.
- It leads to a smaller relative error and a flatter error curve with respect to the smoother applied alone.
- It is fast and usually it obtains a good restored image also without resorting to normal equations.
- It can be combined with other techniques and it can lead to several generalizations (e.g., nonnegativity constraints).

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.

Future work

Theoretical

- A complete theoretical analysis of the regularization properties.

Applications:

- strictly nonsymmetric PSFs.
- Combination with techniques for edge enhancing (Wavelet, Total Variation, . . .).

Numerics/Simulations:

- A complete experimentation with all the proposed BCs (multigrid methods already exist for the arising matrices, see [Aricò, Donatelli, Serra Capizzano, SIMAX, Vol. 26–1 pp. 186–214.]).