Approximated nonstationary iterated Tikhonov with application to image deblurring

MARCO DONATELLI

Dipartimento di Scienza e Alta Tecnologia
Università dell'Insubria - Como

Collaborator: *M. Hanke (U. Mainz, Germany)*

"Congresso SIMAI 2012"
Torino, 25–28 June

Main Issues

1. Image deblurring and regularization
2. Nonstationary Iterated Tikhonov
3. Approximated algorithms
4. Numerical results

Image deblurring problems

$$y = T * x + e$$

- T is large and severely ill-conditioned
- y known, measured data
- e (white Gaussian) noise, s.t. $\|e\| = \delta$

Goal: compute approximation of the noise free solution x

Tikhonov regularization

$T^\dagger y$ is not a good approximation of x because $\|T^\dagger e\|$ is large $\Rightarrow$ regularization.

Tikhonov regularization

Balance the the data fitting and the "explosion" of the solution

$$\min_x \{\|Tx - y\|_2^2 + \alpha\|x\|_2^2\}$$

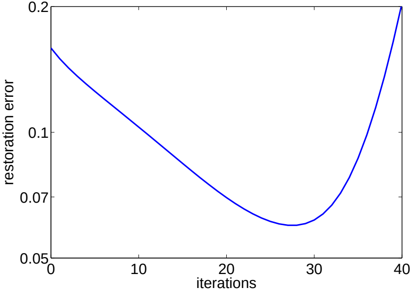
which is equivalent to

$$x = (T^* T + \alpha I)^{-1} T^* y,$$

where $\alpha > 0$ is a regularization parameter.

Iterative regularization methods (semi-convergence)

- Some classical iterative methods firstly reduce the algebraic error into the low frequencies (well-conditioned subspace), when they arrive to reduce the algebraic error into the high frequencies then the restoration error increases because of the noise.
- The regularization parameter is the stopping iteration.



Nonstationary Iterated Tikhonov

Nonstationary iterated Tikhonov is given by

$$x_n = x_{n-1} + T^*(TT^* + \alpha_n I)^{-1}(y - Tx_{n-1}), \tag{1}$$

which is equivalent to

$$x_n = x_{n-1} + (T^*T + \alpha_n I)^{-1}T^*(y - Tx_{n-1}). \tag{2}$$

The iteration (2) can be interpreted as an iterative refinement. Let $e_{n-1} = x - x_{n-1}$ be the error at the step $n-1$, where x is the true solution. Solving the error equation $Te_{n-1} = r_{n-1}$ by Tikhonov, where $r_{n-1} = y - Tx_{n-1}$ is the residual, and using the solution to refine the previous approximation x_{n-1}, we obtain the equation (2).

Convergence

In the noise free case ($\delta = 0$).

Theorem

The method (1) *converges to* $x^\dagger = T^\dagger y$ *if and only if*

$$\lim_{n\to\infty} \sum_{j=1}^{n} \alpha_j^{-1} = \infty.$$

- In the stationary case, $\alpha_n = \alpha$, $\forall n \in \mathbb{N}$, assuming that $x^\dagger = (T^*T)^\nu w$ for some $\nu > 0$ with some $w \in \mathcal{D}((T^*T)^\nu)$, we have
$$\|x_n - x^\dagger\| = \mathcal{O}(n^{-(\nu+1)}).$$

The geometric sequence

- A classical choice for α_n is the successive geometric value

$$\alpha_n = \alpha q^{n-1}, \qquad 0 < q < 1. \tag{3}$$

- Under standard regularity assumptions on $x^\dagger$

$$\|x_n - x^\dagger\| = \mathcal{O}(q^{\nu n}).$$

[Hanke and Groetsh, JOTA, 1998].

The noise case

- Using the discrepancy principle, i.e., the iterative method is stopped at the first value of $n = n(\delta) \geq 1$ for which

$$\|y^\delta - Tx^\delta_{n(\delta)}\| \leq \tau\delta, \tag{4}$$

with $\tau > 1$ fixed.

- For the geometric sequence

$$n(\delta) \leq \mathcal{O}(|\log\delta|).$$

- For the stationary sequence

$$n(\delta) = \mathcal{O}(\delta^{-\frac{2}{2\nu+1}}).$$

Geometric vs stationary sequence

- Under standard regularity assumptions on $x^\dagger$, iterated Tikhonov with the decreasing geometric sequence converges faster than the stationary method both for perturbed and nonperturbed data.
- On the other hand, the main drawback of the decreasing geometric sequence is a steep error curve. Therefore, we need a fair stopping criteria.
- Other nonstationary sequences could be considered if satisfy

$$\lim_{n\to\infty} \sum_{j=1}^{n} \alpha_j^{-1} = \infty.$$

[D., Numer. Algorithms, 2012]

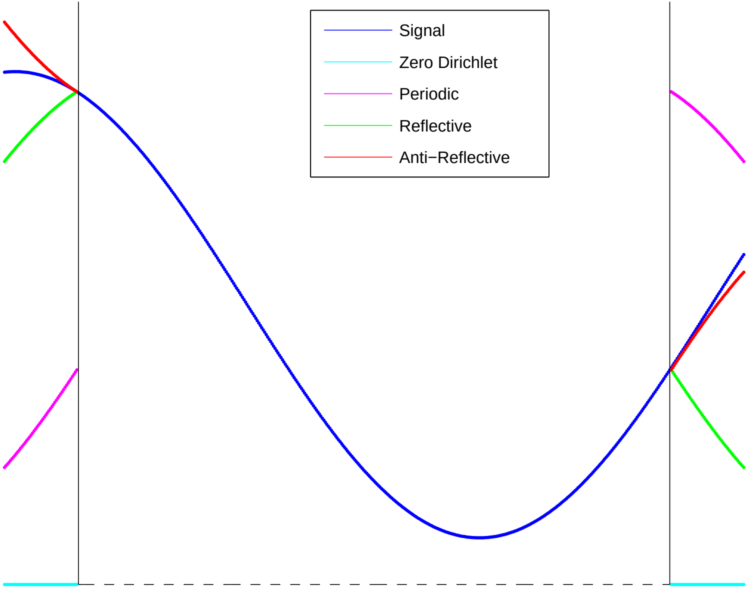
Approximated iterated Tikhonov

- In some applications $TT^* + \alpha_n I$ and $T^*T + \alpha_n I$ are computationally too expensive to invert.
- It is available a good approximation easy to invert.
- Image deblurring with space invariant point spread function:

$$T = P + R + N$$
$$P = \text{BCCB (block circulant circulant block)}$$
$$R = \text{low rank}$$
$$N = \text{low norm.}$$

A good approximation of T can be obtained by P diagonalizzable by 2D discrete Fourier transform.

Boundary conditions for image restoration



Algorithm 1

- Replacing $(TT^* + \alpha_n I)^{-1}$ with $(PP^* + \alpha_n I)^{-1}$ in (1) we obtain

$$x_n = x_{n-1} + T^*(PP^* + \alpha_n I)^{-1}(y - Tx_{n-1}). \qquad (5)$$

- Defining the function

$$\Phi(x) = \frac{1}{2}\|Tx - y\|^2_{(PP^*+\alpha I)^{-1}}$$

in the stationary case $(\alpha_n = \alpha)$ the iteration (5) can be rewritten as

$$x_n = x_{n-1} - \nabla\Phi(x_{n-1}).$$

Algorithm 2 (Nonstationary preconditioning)

- Replacing $(TT^* + \alpha_n I)^{-1}$ with $(PP^* + \alpha_n I)^{-1}$ in (2) we obtain

$$x_n = x_{n-1} + (P^* P + \alpha_n I)^{-1} T^*(y - Tx_{n-1})$$

- In the stationary case, the previous method is the Preconditioned Landweber method previously investigated in [Brianzi et al., SISC, 2008]
- It is a quasi-Newton method for the minimum problem

$$\min_x \frac{1}{2} \| Tx - y \|^2,$$

where the inverse of the Hessian (i.e., $(T^* T)^{-1}$) is replaced with $(P^* P + \alpha_n I)^{-1}$.

Algorithm 2 (Step length)

- The convergence of (14) is not ensured.
- We add a step length τ_n obtaining

$$x_n = x_{n-1} + \tau_n S_n g_{n-1}, \tag{6}$$

where $S_n = (P^* P + \alpha_n I)^{-1}$ and $g_{n-1} = T^*(y - T x_{n-1})$.

- A simple choice for τ_n is

$$\tau_n = \frac{\|g_{n-1}\|^2_{S_n}}{\|g_{n-1}\|^2_{S_n T^* T S_n}}. \tag{7}$$

- It is a simple modification of the steepest descent that is sometimes refereed to as the method of deflected gradients.
- Computing recursively the residual only one matrix-vector product with T and T^* are required.

Algorithm 3 (Periodic refinement)

- Using the iterative refinement interpretation, an approximation of (2) can be derived solving the error equation

$$Pe_n = r_n$$

instead of $Te_n = r_n$, where the residual r_n is defined by the "true" model as $r_n = y - Tx_n$. This produces the iteration

$$x_n = x_{n-1} + (P^*P + \alpha_n I)^{-1} P^*(y - Tx_{n-1}).$$

- For the error equation accurate boundary conditions are not necessary since the error is about uniformly distributed on the domain if the model represented by T is accurate. Hence periodic boundary conditions, such that P is diagonalized by FFT, are an accurate model for the error equation even if they are not good for the original linear system.

Computational cost

- Algorithms 1 and 2 require at each iteration a product with T and T^*, while the Algorithm 3 needs only one product with T.
- Let P be BCCB, then the matrix-vector product with $(P^*P + \alpha_n I)^{-1}P^*$ requires 2 FFTs like $(P^*P + \alpha_n I)^{-1}$.

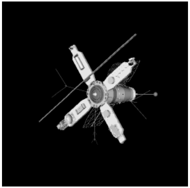
Numerical Experiments

- P is defined by imposing periodic BCs.
- $\alpha_n = \alpha q^{n-1}$ with $\alpha = 1$.
- The relative restoration error (RRE) is

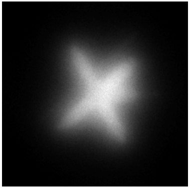
$$\mathrm{RRE} = \frac{\|\tilde{x} - x\|}{\|x\|},$$

where $\tilde{x}$ is the computed solution.

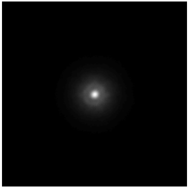
Satelitte test



True

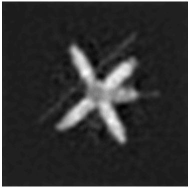


Observed

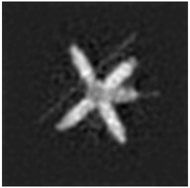


PSF

Optimum restorations ($q = 0.8$)



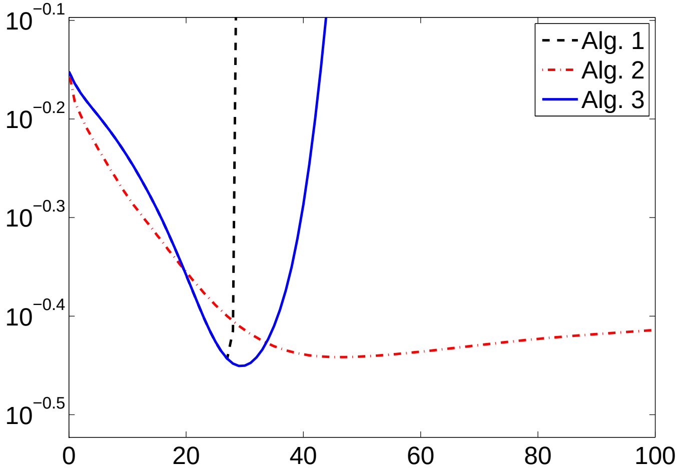
Alg. 1
RRE $= 0.360$
it. 27



Alg. 2
RRE $= 0.362$
it. 47

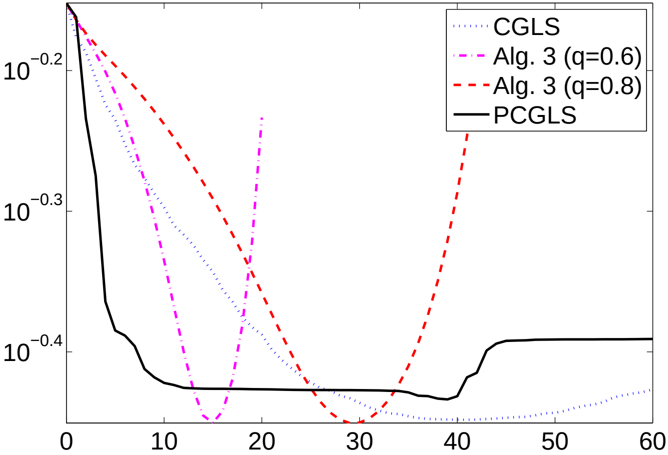
Alg. 3
RRE $= 0.354$
it. 29

RRE varying the iteration number



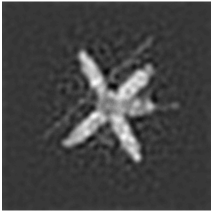
Comparison with CGLS and PCGLS

RestoreTools (Jim Nagy et al.)

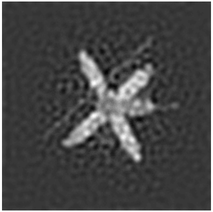


Restorations and RREs

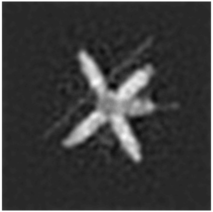
CGLS: 0.35626 − it.:40



PCGLS: 0.36829 − it.:39



Alg. 3 (q=0.6): 0.35446 − it.:15



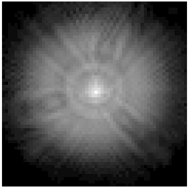
Alg. 3 (q=0.8): 0.35438 − it.:29

Example 2

- Antireflective BCs
- 0.1% of white Gaussian noise
- $A^T \to A'$ (reblurring)

True

Observed



$\log(\mathrm{PSF})$

Restorations and RREs

First row $\rightarrow q = 0.98$, second row $\rightarrow q = 0.9$

Alg. 2 $\tau_n = 1$ (Alg.1) | Algorithm 2 | Algorithm 3

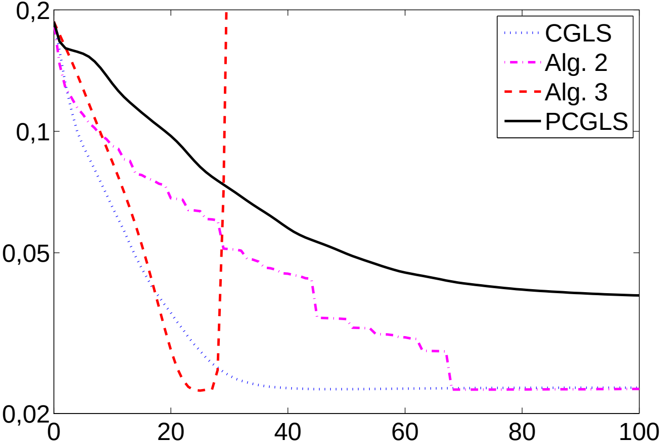


RRE $= 0.0379$, it. 77 | RRE $= 0.0216$, it. 110 | RRE $= 0.0218$, it. 145

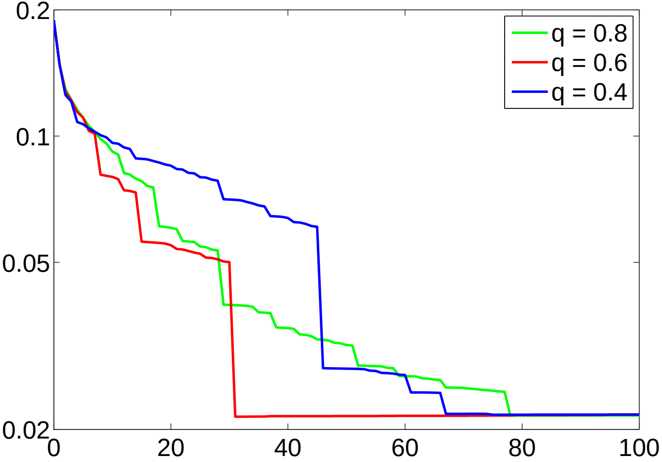


RRE $= 0.072$, it. 18 | RRE $= 0.0213$, it. 95 | RRE $= 0.0214$, it. 43

Comparison with CGLS ($q = 0.8$)



Algorithm 2



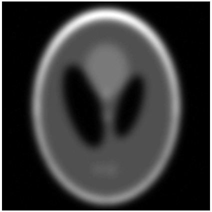
Conclusions and open questions

- How to choose a proper α_n in Algorithms 1 and 2 to ensuring convergence without the step length?
- Algorithm 3
  - Convergence analysis?
  - For a quasi-symmetric PSF it computes a good restoration and it is robust varying the regularization parameter.
  - Accelerates the convergence, like regularizing preconditioning, without spoiling the restoration.
- Other choices of α_n for obtaining a stable convergence?

Example 3 (Nonnegativity constraint)

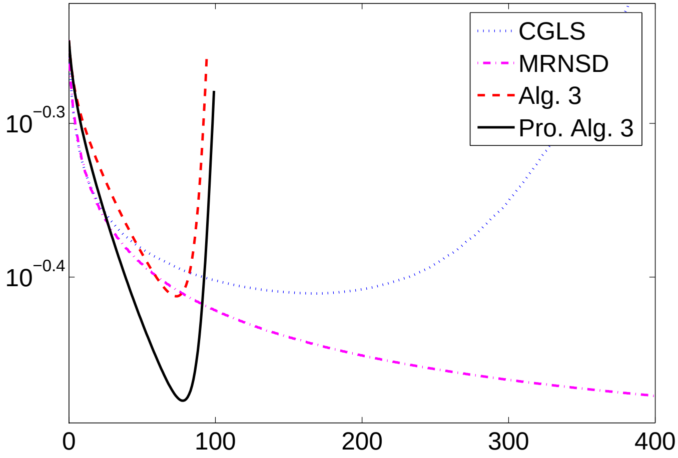
- Gaussian blur
- 1% of white Gaussian noise
- Zero BCs
- Projection of each iteration into the nonnegative cone

True



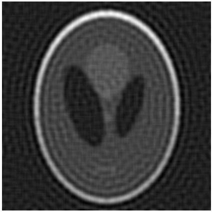
Observed

RRE varying the iteration number ($q = 0.8$)

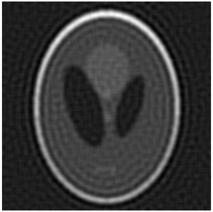


Optimum restorations ($q = 0.8$)

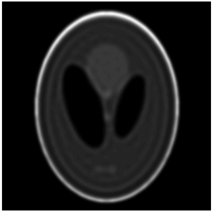
CGLS: 0.38845 − it.:167



Alg. 3: 0.38681 − it.:73



MRNSD: 0.33323 − it.:400



Pro. Alg. 3: 0.33076 − it.:78