Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring

Marco Donatelli

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

Join work with *Martin Hanke*

SIAM IS14
Hong Kong, 12-14 May 2014

Outline

Iterative solution of ill-posed equations

A nonstationary preconditioned iteration

Convergence analysis of the nonstationary iteration

Application to image deblurring

Outline

Iterative solution of ill-posed equations

A nonstationary preconditioned iteration

Convergence analysis of the nonstationary iteration

Application to image deblurring

The model problem

Consider the iterative solution of ill-posed equations

$$Tx = y\,, \tag{1}$$

where $T : \mathcal{X} \to \mathcal{Y}$ is a linear operator between Hilbert spaces. Assume that problem (1) has a solution $x^\dagger$ of minimal norm.

Goal

Compute an approximation of $x^\dagger$ starting from approximate data $y^\delta \in \mathcal{Y}$, instead of the exact data $y \in \mathcal{Y}$, with

$$\|y^\delta - y\| \leq \delta\,, \tag{2}$$

where $\delta \geq 0$ is the corresponding noise level.

Iterative regularization methods

Solving $Tx = y^\delta$ by iterative regularization methods we usually observe one of the following two shortcomings:

- Extremely slow convergence, like Landweber.
- Reasonably fast but may deteriorate if not terminated appropriately (semi-convergence), like conjugate gradient(CGLS).

Preconditioning

Preconditioners can be used to accelerate the convergence, but an imprudent choice of preconditioner may spoil the achievable quality of computed restorations.

Nonstationary iterated Tikhonov regularization

Given x_0 compute for $n = 0, 1, 2, \ldots$

$$z_n = T^*(TT^* + \alpha_n I)^{-1} r_n \,, \qquad r_n = y^\delta - Tx_n \,, \tag{3a}$$

$$x_{n+1} = x_n + z_n \,. \tag{3b}$$

(3a) is equivalent to minimizing the Tikhonov functional

$$\| Tz_n - r_n \|^2 + \alpha_n \| z_n \|^2 \longrightarrow \text{min.} \tag{4}$$

Choices of α_n:

- $\alpha_n = \alpha$, $\forall n$, stationary.
- $\alpha_n = \alpha q^n$ where $\alpha > 0$ and $0 < q \leq 1$, geometric sequence (fastest convergence)

Outline

Iterative solution of ill-posed equations

A nonstationary preconditioned iteration

Convergence analysis of the nonstationary iteration

Application to image deblurring

The starting idea

When linear equations with $T^*T + \alpha I$ or $TT^* + \alpha I$, $\alpha \in \mathbb{R}$, are too expensive to solve we suggest to replace z_n with

$$h_n = C^*(CC^* + \alpha_n I)^{-1} r_n \,, \qquad r_n = y^\delta - Tx_n \,, \tag{5}$$

$$x_{n+1} = x_n + h_n \,, \tag{6}$$

where C is an appropriate approximation of T, i.e., to solve

$$\|Ch_n - r_n\|^2 + \alpha_n \|h_n\|^2 \longrightarrow \text{min.} \tag{7}$$

- This scheme does not require T^* that could be unavailable or expensive to compute in some applications.
- With an appropriate choice of α_n the new scheme turns out to be even faster and more stable than CGLS.

Nonstationary preconditioning

- $P = C^*(CC^* + \alpha_n I)^{-1}$ in (5) would be called preconditioner for the problem $Tx = y^\delta$.
- The iterated Tikhonov method is an iterative refinement procedure applied to the error equation $Te_n \approx r_n$, which is correct up to noise. Hence one may as well consider instead

$$Ce_n \approx r_n \,, \tag{8}$$

  possibly tolerating a slightly larger misfit.
- With this point of view we choose α_n such that the model equation (8) is only solved up to a certain relative amount:

$$\|r_n - Ch_n\| = q_n \|r_n\| \,, \tag{9}$$

  where $q_n < 1$, but not too small.

Choice of q_n and approximation of T

- Assumption: to derive the parameter q_n in (9) we impose

$$\|(C - T)z\| \leq \rho \|Tz\|, \qquad z \in \mathcal{X}, \tag{10}$$

  for some $0 < \rho < 1/2$.
- The condition (10) may be hard to satisfy for a specific problem, as it implies that C and T are spectrally equivalent.

Proposition

Assume that (10) *is satisfied for some* $0 < \rho < 1/2$, *and let* $\tau_* = (1+\rho)/(1-2\rho)$. *Then, if* $\tau_n = \|r_n\|/\delta > \tau_*$, *it follows that*

$$\|r_n - Ce_n\| \leq \left(\rho + \frac{1+\rho}{\tau_n}\right)\|r_n\| < (1-\rho)\|r_n\|. \tag{11}$$

This provides a justification of (8) if $\|r_n\|$ is not too close to δ.

The Algorithm

Algorithm

Let $x_0 \in \mathcal{X}$ *be given, and set* $r_0 = y^\delta - Tx_0$. *Choose* $\tau = (1+2\rho)/(1-2\rho)$ *with* ρ *from* (10), *and fix* $q \in (2\rho, 1)$. *While* $\|r_n\| > \tau\delta$, *let* $\tau_n = \|r_n\|/\delta$, *and compute*

$$h_n = C^*(CC^* + \alpha_n I)^{-1} r_n \,, \tag{12a}$$

where α_n *is such that*

$$\|r_n - Ch_n\| = q_n \|r_n\| \,, \qquad q_n = \max\{\, q,\, 2\rho + (1+\rho)/\tau_n \,\} \,, \tag{12b}$$

and update

$$x_{n+1} = x_n + h_n \,, \qquad r_{n+1} = y^\delta - Tx_{n+1} \,. \tag{12c}$$

Details of Algorithm

- The parameter q is meant as a safeguard to prevent that the residual decreases too rapidly.
- If ρ happens to be too small, or if (10) is only satisfied approximately, then error components may build up strongly;
- There is a unique positive regularization parameter α_n that determines h_n in the prescribed manner and it can be computed with few steps of a Newton scheme.
- The stopping rule is the discrepancy principle: terminate after $n = n_\delta \geq 0$ iterations with

$$\|r_{n_\delta}\| \leq \tau\delta < \|r_n\|, \qquad n = 0, 1, \ldots, n_\delta - 1. \tag{13}$$

Outline

Iterative solution of ill-posed equations

A nonstationary preconditioned iteration

Convergence analysis of the nonstationary iteration

Application to image deblurring

Theoretical properties of Algorithm

Proposition

Under assumption (10) *the norm of the iteration error* $e_n = x^\dagger - x_n$ *of Algorithm decreases monotonically for* $n = 0, 1, \ldots, n_\delta - 1$.

Corollary

With the same assumption and notation there holds

$$\|e_0\|^2 \geq c \sum_{n=0}^{n_\delta - 1} \|r_n\|^2 \tag{14}$$

for some constant $c > 0$, *depending only on* ρ *and* q.

$\Rightarrow$ Algorithm terminates after finitely many iterations, when $\delta > 0$.

Convergence

Theorem (Noise free)

Assume that the data are exact, i.e., $\delta = 0$*, and that* x_0 *is no solution of problem* (1)*. Then the sequence* $(x_n)_n$ *converges as* $n \to \infty$ *to the solution of* (1) *that is closest to* x_0*.*

Theorem (Regularization)

Assume that (10) *holds for some* $0 < \rho < 1/2$*, and let* $\|y^\delta - y\| \leq \delta$ *be true for all* $\delta > 0$*. For fixed parameters* τ *and* q *denote by* x^δ *the approximations computed by the Algorithm. Then, as* $\delta \to 0$*,* x^δ *converges to the solution of* (1) *that is closest to* x_0 *in the norm of* $\mathcal{X}$*.*

Outline

Iterative solution of ill-posed equations

A nonstationary preconditioned iteration

Convergence analysis of the nonstationary iteration

Application to image deblurring

Image deblurring problems

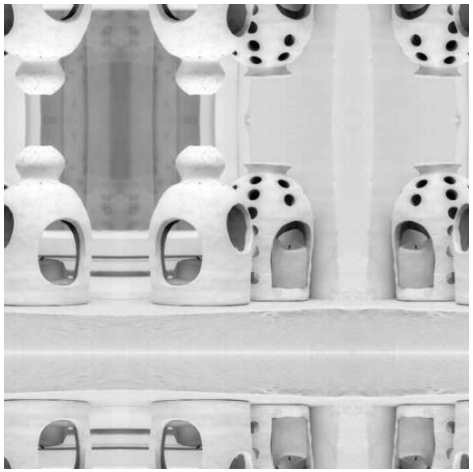
$\mathbf{y}^\delta \quad = \quad \mathbf{T} \quad * \quad \mathbf{x} \quad + \quad \mathbf{e}$

- $\mathbf{T}$ is large and severely ill-conditioned (discretizzation of a Fredholm integral equations of the first kind)
- $\mathbf{y}^\delta$ known, measured data (blurred and noisy image)
- $\mathbf{e}$ (white Gaussian) noise, s.t. $\|e\| = \delta$

Boundary Conditions (BCs)

zero Dirichlet

Periodic



Reflective

Antireflective

The matrix $\mathbf{C}$

Space invariant point spread function (PSF)

$$\Downarrow$$

$\mathbf{T}$ has a doubly Toeplitz-like structure that carries the "correct" boundary conditions.

- doubly circulant matrix $\mathbf{C}$ diagonalizzable by FFT, that corresponds to periodic BCs.
- The boundary conditions have a very local effect

$$\mathbf{T} - \mathbf{C} = \mathbf{E} + \mathbf{R}, \tag{15}$$

where $\mathbf{E}$ is of small norm and $\mathbf{R}$ of small rank.

Choice of parameters

- We do not know whether a rigorous estimate of

$$\|(\mathbf{T} - \mathbf{C})\mathbf{z}\| \leq \rho \, \|\mathbf{T}\mathbf{z}\|$$

  will hold for all (relevant) vectors $\mathbf{z}$ and $\rho < 1/2$.
  According to our numerical tests, the estimate (11) is satisfied with ρ of a few percent, i.e., 10^{-3} or 10^{-2}.
- $q = 0.7$, but all $q \in [0.6,\ 0.8]$ gives comparable results.

Other methods:

- CGLS, P-CGLS, our algorithm with $\alpha_n = 0.5\, q^n$ ($\|\mathbf{T}\|_\infty \approx 1$).
- Stopping by discrepancy principle with $\tau = 1.01$.

Numerical results

- We add to the blurred image a white Gaussian noise with the relative amount of noise

$$\nu = \frac{\delta}{\|\mathbf{y}\|}$$

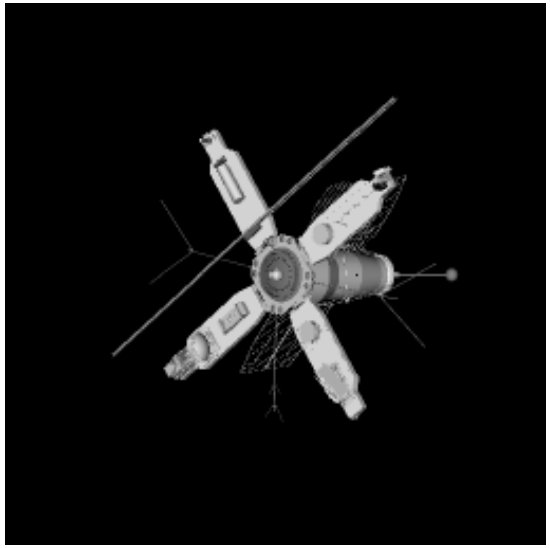
- To compare the quality of the restorations, we evaluate their relative restoration errors (RRE), i.e.,

$$\text{RRE} = \frac{\|\mathbf{x} - \mathbf{x}^\dagger\|}{\|\mathbf{x}^\dagger\|},$$

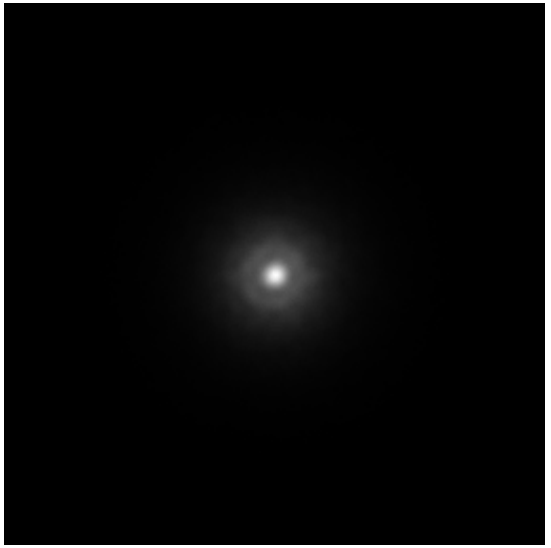
  where $\mathbf{x}$ is the computed solution.
- CGLS and P-CGLS from the Matlab Toolbox RestoreTools
  J. Nagy et al. http://www.mathcs.emory.edu/~nagy/RestoreTools

Example 1

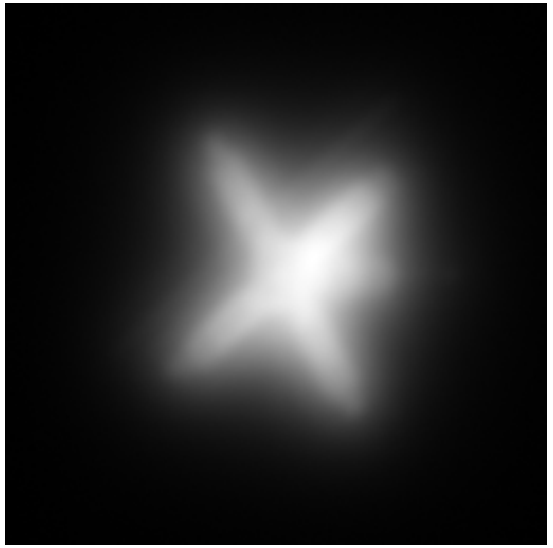
Size 256×256, $\nu = 0.5\,\%$, $\rho = 10^{-3}$, zero Dirichlet BCs.



True image

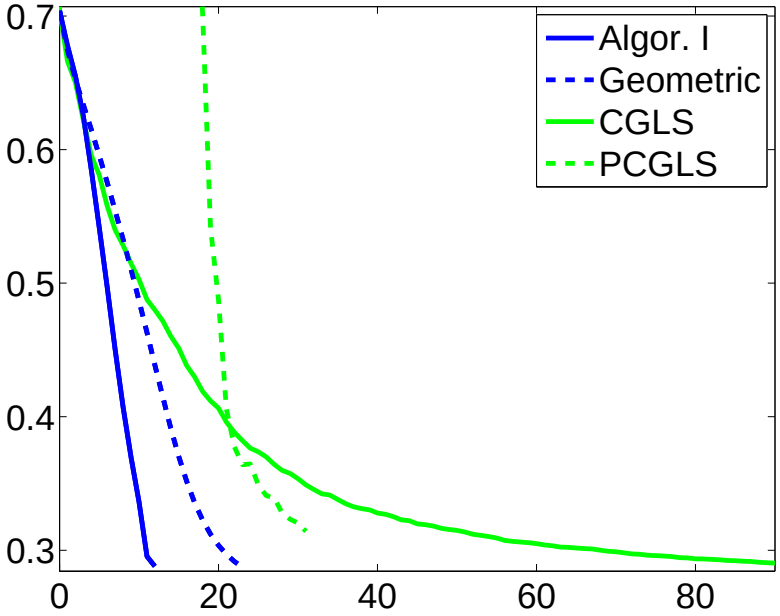


PSF

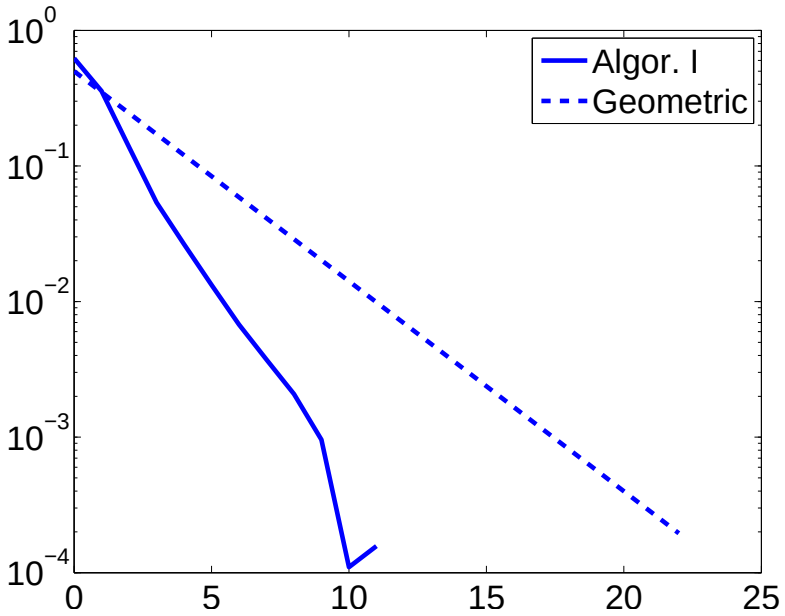


Observed image

RRE and α_n

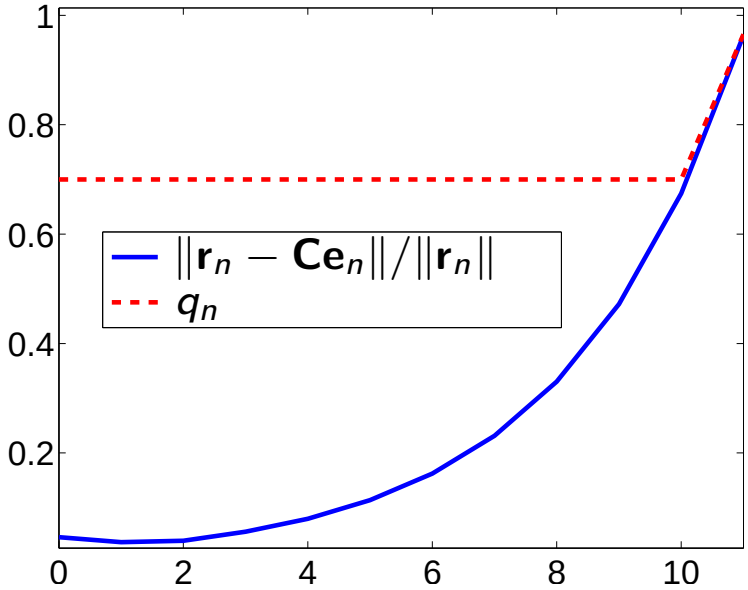


RRE vs iterations

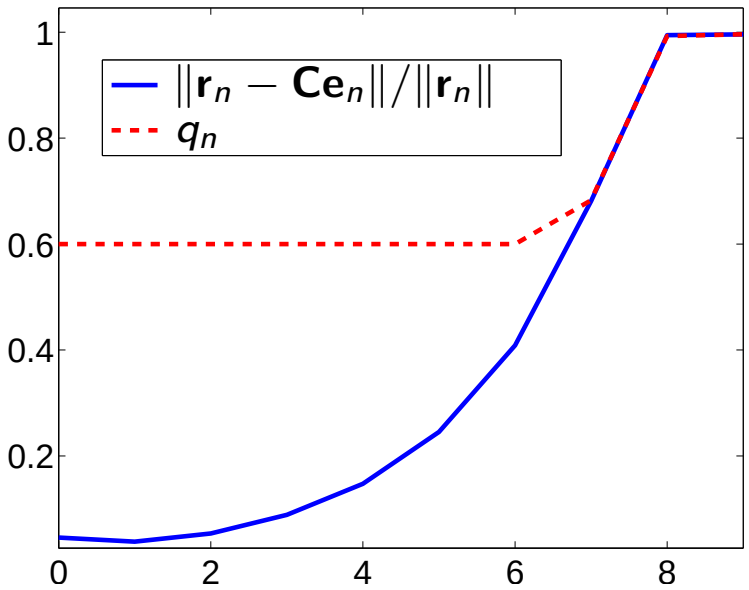


α_n vs n

Check of the assumption



$q = 0.7$

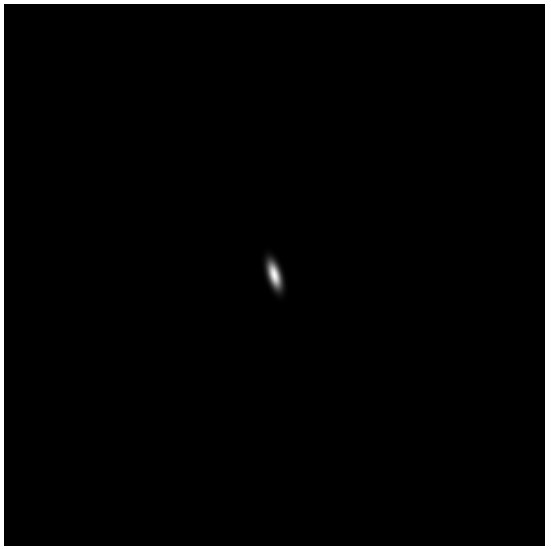


$q = 0.6$

Example 2

Size 237×237, $\nu = 0.1\,\%$, $\rho = 10^{-3}$, zero Dirichlet BCs.

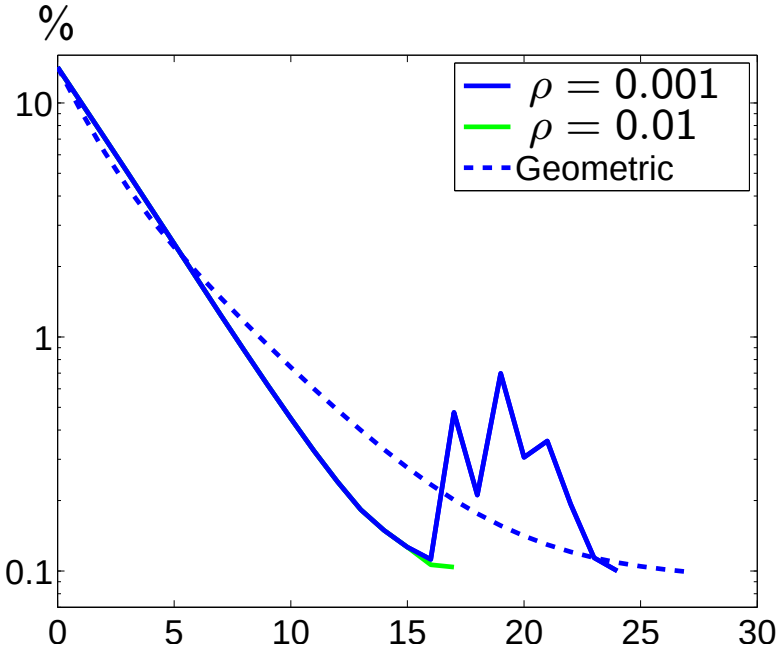
True image

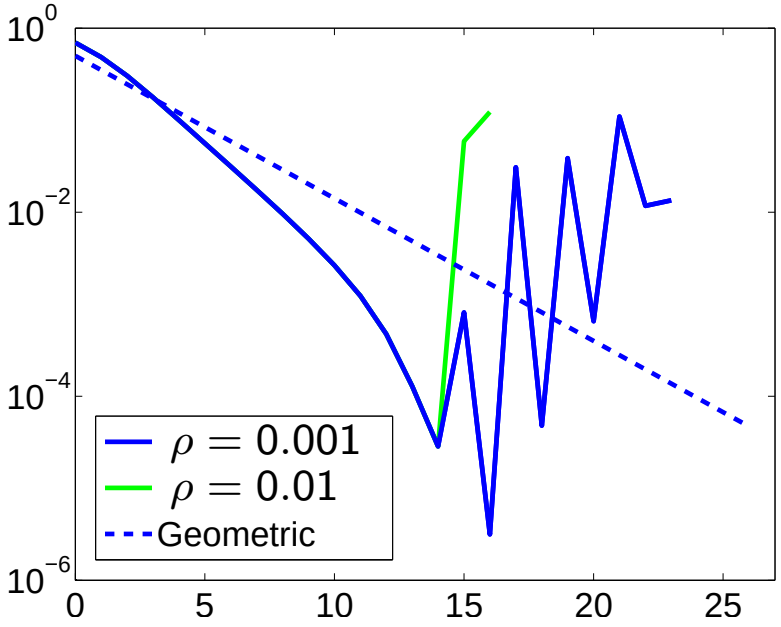


PSF

Observed image

Detect that $\rho = 10^{-3}$ is too small for the assumption



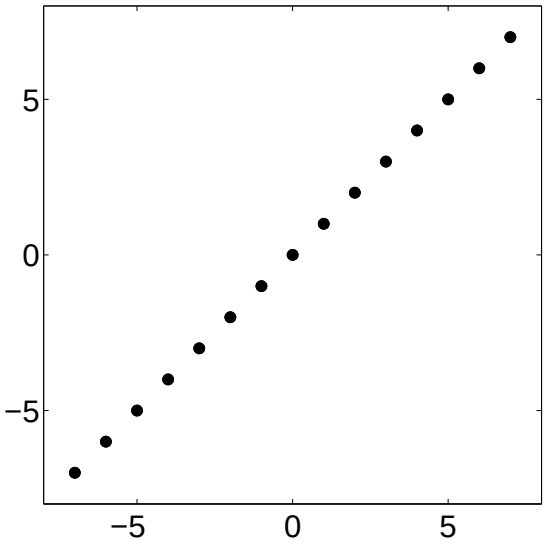


- The algorithm often recovers (RRE $= 0.254$).
- Increasing $\rho = 10^{-2} \Rightarrow$ the zigzagging disappears and the Algorithm terminates with RRE $= 0.261$.
- Different ρ only affects the final stage of the iteration.

Example 3

- Size 452×452, $\nu = 1\,\%$, Antireflective BCs.
- The structure of $\mathbf{T}$ is more involved $\Rightarrow \rho = 10^{-2}$

True image

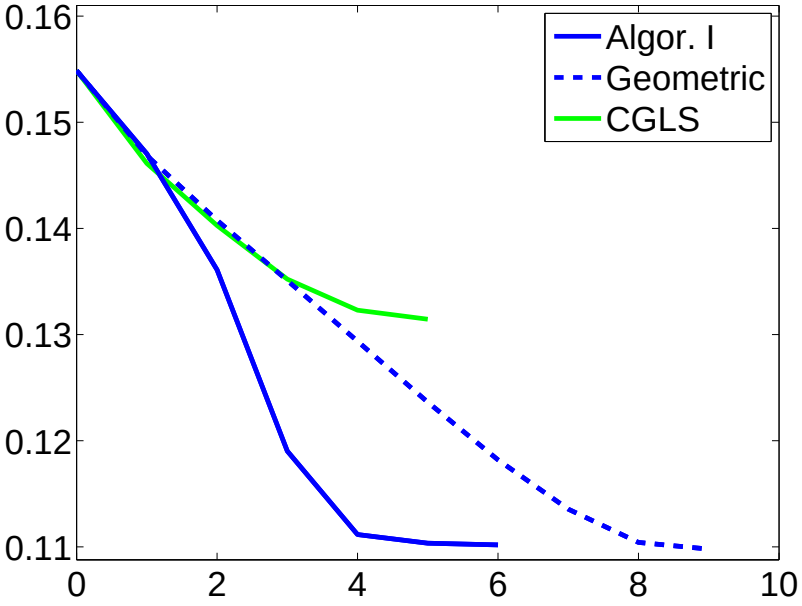


PSF

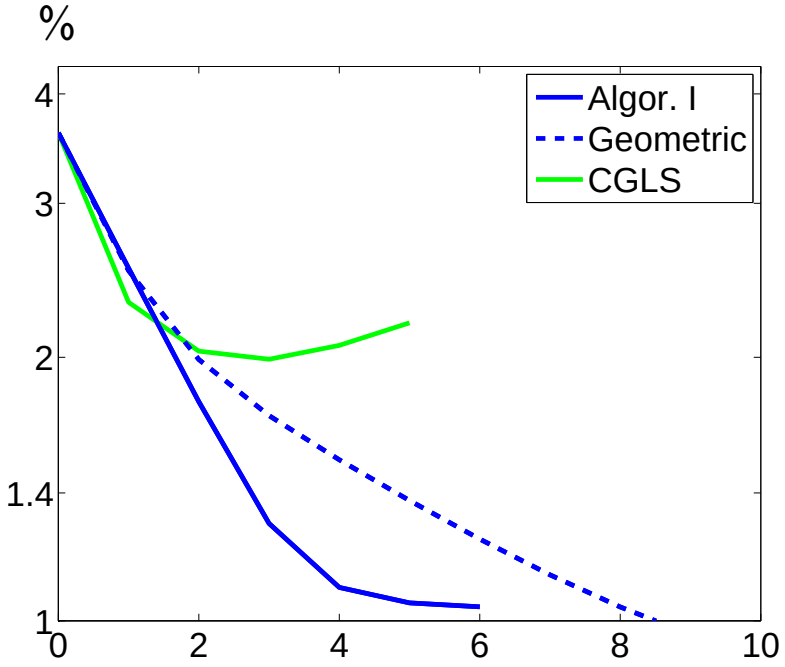
Observed image

Avoid $\mathbf{T}^*$

- Restorations in the range of $\mathbf{T}^*$ come with unwanted boundary artefacts [D., Serra-Capizzano IP 2005].
- RestoreTools implements the rebluring strategy: replaces $\mathbf{T}^*$ with $\mathbf{T}'$ by PSF rotated by 180°, but $\mathbf{T}'\mathbf{T}$ is not symmetric.



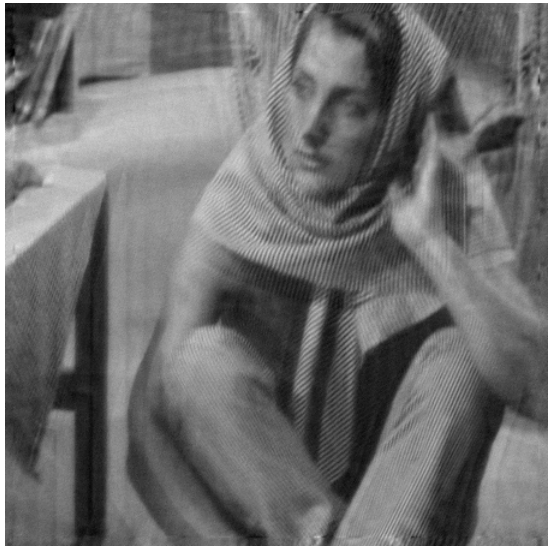
RRE vs iterations



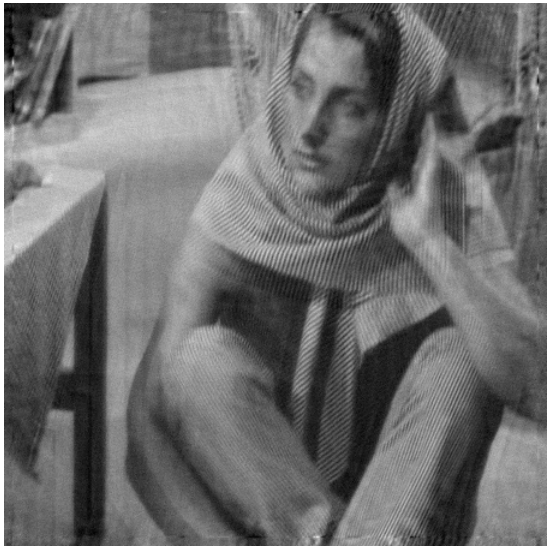
$\|\mathbf{r}_n\|/\|\mathbf{y}^\delta\|$ vs n

Restored images

CGLS stopped at the minimum RRE.



Our Algorithm
RRE = 0.110, it. 6



Geometric Sequence
RRE = 0.110, it. 9

CGLS
RRE = 0.131, it. 5

Conclusions

- Under the assumption that an appropriate approximation C of T is available, our new scheme turns out to be fast and stable.
- The choice of ρ reflect how much we trust in the previous approximation.
- A too small choice of ρ can be detected by α_n or $\|r_n\|$.
- The choice of q is not crucial ($q = 0.7$ balance between speed and stability).

References

M. Donatelli and M. Hanke
Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring,
Inverse Problems, 29 (2013) 095008.

The choice of α_n is inspired by

M. Hanke
A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems,
Inverse Problems, 13 (1997), pp. 79–95.