Regularization preconditioners for frame-based deconvolution

MARCO DONATELLI

Dept. of Science and High Tecnology – U. Insubria (Italy)

Joint work with
M. Hanke (U. Mainz), *D. Bianchi* (U. Insubria),
Y. Cai, T. Z. Huang (UESTC, P. R. China)

OIP 2016 - 19-21 September 2016

Outline

Outline

The model problem

Consider the solution of ill-posed equations

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

where $T : \mathcal{X} \to \mathcal{Y}$ is a linear operator between Hilbert spaces.

- T is a compact operator, the singular values of T decay gradually to zero without a significant gap.
- 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

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

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

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

- T is doubly Toeplitz, large and severely ill-conditioned (discretizzation of an integral equations of the first kind)
- y^{δ} are known measured data (blurred and noisy image)
- ξ is noise; $\|\xi\| = \delta$

$\longrightarrow$ discrete ill-posed problems (Hansen, 90's)

- The singular values of T are large in the low frequencies, decays rapidly to zero and are small in the high frequencies.
- The solution of $Tx = y^\delta$ requires some sort of regularization:

$$x = T^\dagger y^\delta = x^\dagger + T^\dagger \xi,$$

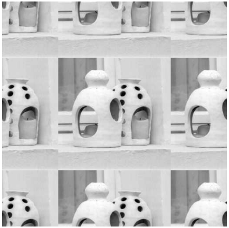
  where $\|T^\dagger \xi\|$ is large.
- Tikhonov regularization: balance the data fitting and the "explosion" of the solution

$$\min_x \{\|Tx - y^\delta\|^2 + \alpha\|x\|^2\} \iff x = (T^*T + \alpha I)^{-1} T^* y^\delta$$

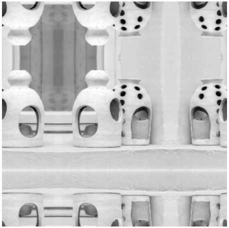
  where $\alpha > 0$ is a regularization parameter.

Boundary Conditions (BCs)

zero Dirichlet



Periodic



Reflective

Antireflective

The matrix C

Space invariant point spread function (PSF)

⇓

T has a doubly Toeplitz-like structure that carries the "correct" boundary conditions.

- doubly circulant matrix C diagonalizable by FFT, that corresponds to periodic BCs.
- The boundary conditions have a very local effect

$$T - C = E + R, \quad (3)$$

where E is of small norm and R of small rank.

Outline

Synthesis approach

- Images have a sparse representation in the wavelet domain.
- Let W^* be a wavelet or tight-frame synthesis operator ($W^*W = I$) and v the frame coefficients such that

$$x = W^* v.$$

- A regularized solution can be obtained in terms of the wavelet coefficients by

$$\min_v \{\|Kv - y^\delta\|^2 + 2\mu\|v\|_1\}$$

where $K = TW^*$.

Iterative soft-thresholding algorithm (ISTA)

- The solution of

$$\min_v \{ \|Kv - y^\delta\|^2 + 2\mu \|v\|_1 \}$$

can be computed by Iterative soft-thresholding algorithm (ISTA) [Daubechies, Defrise, De Mol, CPAM 2004]

$$\begin{cases} z^{n+1} = v^n + K^*(y^\delta - Kv^n), \\ v^{n+1} = S_\mu(z^{n+1}), \end{cases} \tag{4}$$

when $\|K\| < 1$ and where S_μ denotes the soft-thresholding .

- The inner step is a Landweber iteration

- The Linearized Bregman splitting proposed in [Yin, Osher, Goldfarb, Darbon, SIIMS 2008]

$$\begin{cases} z^{n+1} = z^n + K^*(y^\delta - Kv^n), \\ v^{n+1} = S_\mu(z^{n+1}), \end{cases} \tag{5}$$

  is analogous to ISTA replacing v^n with z^n at the first occurrence in the inner step.
- With $z^0 = v^0 = 0$ converges to the unique minimizer of

$$\min_{v \in \mathbb{R}^s} \left\{ \mu \|v\|_1 + \frac{1}{2}\|v\|^2 : Kv = y^\delta \right\} \tag{6}$$

  if $\|K\| < 1$.
- Common feature: slow convergence

Modified Linearized Bregman algorithm (MLBA)

- The MLBA proposed in [Cai, Osher, Shen, SIIMS 2009]

$$\begin{cases} z^{n+1} = z^n + K^* P(y^\delta - Kv^n), \\ v^{n+1} = S_\mu(z^{n+1}), \end{cases} \tag{7}$$

where $z^0 = v^0 = 0$.

- It is the linearized Bregman splitting applied to the preconditioned equation

$$P^{1/2}Kv = P^{1/2}y^\delta.$$

- $P = (TT^* + \alpha I)^{-1} \Rightarrow$ the iteration (7) converges to the unique minimizer of

$$\min_{v \in \mathbb{R}^s} \left\{ \mu \|v\|_1 + \frac{1}{2}\|v\|^2 : v = \arg\min_{v \in \mathbb{R}^s} \|Kv - y^\delta\|_P^2 \right\} \tag{8}$$

within few iterations.

Choosing P in MLBA

- Similarly, a preconditioned version of ISTA with $P = (TT^* + \alpha_n I)^{-1}$ has been studied in [Huang, D., Chan, IPI 2013]:

$$\begin{cases} z^{n+1} = v^n + K^*(TT^* + \alpha_n I)^{-1}(y^\delta - Kv^n), \\ v^{n+1} = S_\mu(z^{n+1}), \end{cases} \qquad (9)$$

- $K^*(TT^* + \alpha_n I)^{-1} r^n = (K^*K + \alpha_n I)^{-1} K^* r^n \implies$ Tikhonov regularization on the residual $r^n = y^\delta - Kv^n$.
- Possible choice $\alpha_n = \alpha_0 q^n$ with $0 < q < 1$, e.g. $\alpha_n = 0.8^n$.
- The inversion of $TT^* + \alpha_n I$ could be costly $\Rightarrow$ the authors of MLBA proposed the use of $P = (CC^* + \alpha I)^{-1}$

$$\Downarrow$$

it does not provide an accurate restoration when the BCs are essential!

Outline

Preconditioned regularization

Replace the original problem $Tx = y^\delta$ with

$$P^{-1}Tx = P^{-1}y^\delta$$

such that

1. inversion of P is cheap
2. $P \approx T$ but not too much ($T^\dagger$ unbounded while P^{-1} must be bounded!)

Alert!

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

Classical preconditioner

- Historically, the first attempt of this sort was by Hanke, Nagy, and Plemmons (1993): In that work

$$P = C_\varepsilon,$$

  where C_ε is the optimal doubly circulant approximation of T, with eigenvalues set to be one for frequencies above $1/\varepsilon$.
  Very fast, but the choice of ε is delicate and not robust.
- Subsequently, other regularizing preconditioners have been suggested: Bertero and Piana (1997), Kilmer and O'Leary (1999), Estatico (2002), Egger and Neubauer (2005), Brianzi, Di Benedetto, and Estatico (2008), D. and Hanke (2013), Dell'Acqua, D., and Estatico (2014)

- We look for a fast and accurate preconditioned Landweber-like iteration.
- The preconditioner in MLBA is analogous to the preconditioned Landweber in [Brianzi, Di Benedetto, Estatico, SISC 2009] $\Rightarrow$ in order to guarantee the convergence α could be large reducing the acceleration.
- A recent nonsymmetric strategy

$$ZTx = Zy^{\delta},$$

where Z is a regularized approximation of $T^{\dagger}$, solved by the Van Cittert iteration [Dell'Acqua, D., Estatico, JCAM 2014]:

$$x^{n+1} = x^n + Z(y^{\delta} - Tx^n) = x^n + (C^*C + \alpha I)^{-1}C^*(y^{\delta} - Tx^n)$$

$$\Downarrow$$

it can be seen as an approximated iterated Tikhonov.

Iterated Tikhonov

- The correction step

$$h^n = (TT^* + \alpha I)^{-1} T^* (y^\delta - Tx^n)$$

  is the Tikhonov solution of the error equation $Te^n = r^n$, where $r^n = y^\delta - Tx^n$ is the residual.
- The iterative refinement

$$x^{n+1} = x^n + h^n$$

  is called Iterated Tikhonov.
- In the noise free case ($\delta = 0$) we solve by Tikhonov the true error equation $Te^n = y - Tx^n$.
- In the noisy case ($\delta > 0$) the iterative refinement is correct up to the noise level δ since the residual r^n differ from the true residual $y - Tx^n$.

Approximated Iterated Tikhonov (AIT)

- In the noisy case, we could admit a further misfit "lower" than δ:

$$\min_{h^n}\{\|Th^n - r^n\|^2 + \alpha\|h^n\|^2\} \quad \Longrightarrow \quad \min_{h^n}\{\|Ch^n - r^n\|^2 + \alpha\|h^n\|^2\}$$

  which solution is

$$h^n = (C^*C + \alpha I)^{-1}C^*(y^\delta - Tx^n).$$

- $P = (C^*C + \alpha I)^{-1}C^*$ is a nonsymmetric preconditioner which leads to the preconditioned Van Cittert iteration

$$x^{n+1} = x^n + P(y^\delta - Tx^n).$$

- Convergence analysis based on the knowledge of the complex eigenvalues of PT and requiring a complex step length [Dell'Acqua, D., Estatico, JCAM 2014].

Nonstationary AIT

- Nonstationary preconditioning [D., Hanke, IP 2013]
- Use a sequence of regularization parameters $\{\alpha_n\}$ obtaining

$$x^{n+1} = x^n + P_n(y^\delta - Tx^n), \text{ with } P_n = (C^*C + \alpha_n I)^{-1}C^*.$$

- We need the "strong" assumption

$$\|(C - T)u\| \leq \rho \|Tu\|, \qquad \forall u \in \mathcal{X},$$

for some $0 < \rho < 1/2$.

- Compute α_n by few Newton's steps for the equation

$$\|r^n - CPr^n\| = q_n \|r^n\|,$$

where $q_n < 1$ depends on ρ and δ but it is not too small.
If C is diagonalizzable by FFT then the computational cost is linear in the number of pixels.

The Algorithm

Algorithm

Let $x^0 \in \mathcal{X}$ be given, and set $r^0 = y^\delta - Tx^0$.
Choose $\tau = (1+2\rho)/(1-2\rho)$ and fix $q \in (2\rho, 1)$.
While $\|r^n\| > \tau\delta$, let $\tau_n = \|r^n\|/\delta$, and compute α_n s.t.

$$\|r^n - Ch^n\| = q_n\|r^n\|, \qquad q_n = \max\{\, q,\, 2\rho + (1+\rho)/\tau_n \,\}. \tag{10a}$$

Then, update

$$h^n = (C^*C + \alpha_n I)^{-1} C^* r^n, \tag{10b}$$

$$x^{n+1} = x^n + h^n, \qquad r^{n+1} = y^\delta - Tx^{n+1}. \tag{10c}$$

Details

- The parameter q prevents that r_n decreases too rapidly.
- The unique α_n can be computed by Newton iteration.
- The stopping rule is almost the discrepancy principle.

Theoretical results

Theorem

The norm of the iteration error $e^n = x^\dagger - x^n$ decreases monotonically as long as

$$\|r^n\| \leq \tau\delta \leq \|r^{n-1}\|, \qquad \tau > 1 \text{ fixed.}$$

Theorem

For exact data ($\delta = 0$) the iterates x^n converges to the solution of $Tx = y$ that is closest to x^0 in the norm of $\mathcal{X}$.

Theorem

For noisy data ($\delta > 0$), as $\delta \to 0$, the approximation x^δ converges to the solution of $Tx = y$ that is closest to x^0 in the norm of $\mathcal{X}$.

Outline

AIT + linearized Bregman splitting

- Replace preconditioned Landweber with AIT in MLBA.
- Usual assumption

$$\|(C - T)u\| \leq \rho \|Tu\|, \qquad \forall u \in \mathcal{X}$$

- Further assumption

$$\|CW^*(v - S_\mu(v))\| \leq \rho\delta, \qquad \forall v \in \mathbb{R}^s, \tag{11}$$

which is equivalent to consider the soft-threshold parameter $\mu = \mu(\delta)$ and such that $\mu(\delta) \to 0$ as $\delta \to 0$.

AIT + linearized Bregman splitting – 2

Algorithm [Cai, D., Bianchi, Huang, 2016]

$$\begin{cases} z^{n+1} = z^n + WC^*(CC^* + \alpha_n I)^{-1}(y^\delta - TW^*v^n), \\ v^{n+1} = S_\mu(z^{n+1}), \end{cases} \tag{12}$$

where the parameter (α_n, stopping iteration, etc.) are fixed as in AIT.

Theorem

For noisy data ($\delta > 0$), as $\delta \to 0$, the approximation x^δ converges to the solution of $Tx = y$ that is closest to z^0 in the ℓ_2-norm.

In the applications, if an estimation of the best α is available, we can fix $\alpha_n = \alpha_{\rm opt}$, but there is not any convergence result.

Outline

- W by piecewise linear B-spline framelets. In 1D, given the masks

$$b_0 = \frac{1}{4}[1,\ 2,\ 1], \quad b_1 = \frac{\sqrt{2}}{4}[1,\ 0,\ -1], \qquad b_2 = \frac{1}{4}[-1,\ 2,\ -1]$$

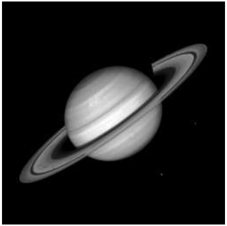
  the associated matrices are defined by imposing reflective BCs. The nine 2D filters are obtained by tensor product.
- $\rho = 10^{-4}$, $q = 0.5$.
- $\mathrm{PSNR} = 20 \log_{10} \frac{255 \cdot n}{\|x - \tilde{x}\|}$, with $\tilde{x}$ the computed approximation.
- Best regularization parameter μ and α when necessary are estimated by trial and error for every method in order to achieve the max PSNR.

Compared methods

- MLBA: iteration (7) by Cai, Osher, and Shen.
- AIT–Breg: our nonstationary iteration (12).
- AIT–Breg–opt: our iteration (12) with a stationary $\alpha_n = \alpha_{\text{opt}}$ chosen by hand like in MLBA.
- FA–MD, TV–MD: ADMM [Almeida, Figueiredo, IEEE 2013] for Frame-based Analysis and Total Variation, respectively (rectangular matrix).
- FTVd: extension of FTVd in [Wang et al. SIIMS 2008] to deal with boundary artifacts [Bai et al., 2014] (enlarging of the domain).

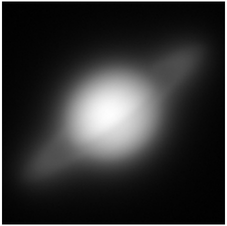
Example 1 (Saturn)

$\nu = 1\,\%$, Zero BCs.



True image

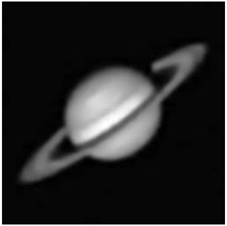
PSF



Observed image

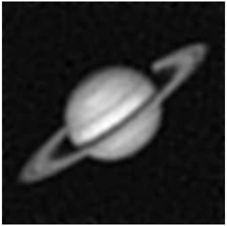
Restorations

Method	PSNR	CPU time
AIT–Breg	31.25	10.32
AIT–Breg–opt	31.49	16.56
MLBA	30.97	200.99
FA–MD	30.87	90.85
TV–MD	31.17	47.61
FTVd:	30.50	1.75



MLBA

AIT–Breg

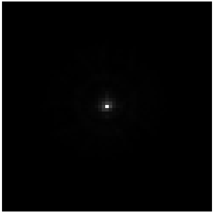


FTVd

Example 2 (Boat)

$\nu = 1\,\%$, Antireflective BCs.

True image



PSF

Observed image

Restorations

Method	PSNR	CPU time
AIT–Breg	29.77	19.57
AIT–Breg–opt	30.17	3.67
MLBA	29.43	34.26
FA–MD	29.61	15.95
TV–MD	29.87	16.74
FTVd:	28.95	0.73

MLBA

AIT–Breg–opt

TV–MD

Conclusions and future work

- Regularization preconditioners can be used to speed up the convergence of nonlinear methods without losing the quality of the restoration.
- Under the assumption that an 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 and in practice it can be small enough.
- Our scheme does not require T^*.
- The proposed regularizing preconditioners could be combined also with ISTA, with the analysis approach, etc.

References

Y. Cai, M. Donatelli, D. Bianchi, T. Z. Huang
Regularization preconditioners for frame-based image deblurring with reduced boundary artifacts,
SIAM J. Sci. Comput., 38–1 (2016), pp. B164–B189.

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

J. Huang, M. Donatelli, R. Chan
Nonstationary Iterated Thresholding Algorithms for Image Deblurring,
Inverse Problems and Imaging, 7–3 (2013), pp. 717–736.