A Preconditioned Version of a Nested Primal-Dual Algorithm for Image Deblurring

Stefano Aleotti [1] Marco Donatelli [1] Rolf Krause [2] Giuseppe Scarlato [1]

[1]Università degli Studi dell'Insubria [2]King Abdullah University of Science and Technology (KAUST)

Aim of the work

This work aims to accelerate the convergence of the *Nested Primal-Dual* (NPD) algorithm [3] by introducing a left preconditioner. We propose a new algorithm, the *Preconditioned Nested Primal-Dual* (PNPD), and prove its convergence to a model where the data fidelity term uses an energy norm induced by the preconditioner instead of the usual euclidean norm. Furthermore, we compare the performance of PNPD with the original NPD algorithm and a variable metric version of the NPD algorithm, called *Nested Primal-Dual Iterated Tikhonov* (NPDIT) [2], which can be seen as a right preconditioned version of NPD.

Motivation

Inverse problems in imaging, such as image deblurring, denoising, and computed tomography, often lead to optimization problems of the form

$$\underset{u\in\mathbb{R}^d}{\operatorname{argmin}}\ \underbrace{f(u)}_{\text{smooth}} + \underbrace{h(Wu)}_{\text{convex}}, \tag{1}$$

where $f:\mathbb{R}^d\to\mathbb{R}$ is a convex and smooth function, $h:\mathbb{R}^{d'}\to\mathbb{R}\cup\{\infty\}$ is convex and possibly non-smooth, and $W\in\mathbb{R}^{d'\times d}$ is a linear operator. Since h is not differentiable, we can not use smooth optimization methods. A common approach to solve this kind of problem is to use proximal methods that rely on the *proximal operator* of the convex function $g = h\circ W$:

$$\operatorname{prox}_{\alpha g}(u) = \underset{v\in\mathbb{R}^d}{\operatorname{argmin}}\ \alpha g(v) + \frac{1}{2}\|u-v\|^2. \tag{2}$$

When the proximal operator can not be computed in closed form, like for the *Total Variation* (TV) regularization, inexact proximal methods can be used. An example of such methods is the *Nested Primal-Dual* (NPD) algorithm, which is a first-order iterative method that alternates a gradient descent step on the smooth part with an inexact proximal step on the convex part, implemented via a nested iteration. Furthermore, to accelerate the convergence of the algorithm, NPD uses a Nesterov-like extrapolation step:

NPD

$$\begin{cases}\bar{u}_n = u_n + \gamma_n(u_n - u_{n-1}), & \leftarrow \text{Nesterov extrap.}\\ u_{n+\frac{1}{2}} = \bar{u}_n - \alpha_n\nabla f(\bar{u}_n), & \leftarrow \text{gradient descent}\\ u_{n+1} \approx \operatorname{prox}_{\alpha_n h\circ W}(u_{n+\frac{1}{2}}). & \leftarrow \text{proximal step}\end{cases} \tag{3}$$

The proximal step is obtained by a dual sequence: Let $u\in\mathbb{R}^d$, $v^0\in\mathbb{R}^{d'}$, $\alpha>0$, $0<\beta<2/\|W\|^2$, and define the sequence

$$v^{k+1} = \operatorname{prox}_{\beta\alpha^{-1}h^*}(v^k + \beta\alpha^{-1}W(u - \alpha W^T v^k)). \tag{4}$$

having limit $\lim_{k\to\infty} v^k = \hat{v}$. Then we have

$$\operatorname{prox}_{\alpha h\circ W}(u) = u - \alpha W^T\hat{v}.$$

Model and Problem

We focus on the image deblurring linear model

$$Au = b^\delta, \tag{5}$$

where $A\in\mathbb{R}^{s\times d}$ represents the discretization of a space invariant convolution operator, $u\in\mathbb{R}^d$ denotes an unknown two-dimensional image with d pixels, $b^\delta\in\mathbb{R}^s$ contains the observed image corrupted by white Gaussian noise η_δ, and $\delta\geq 0$ is the noise level. The ill-posed nature of the operator A and the presence of noise requires a regularization strategy to solve problem (5). A common approach involves solving the optimization problem

$$\underset{u\in\mathbb{R}^d}{\operatorname{argmin}}\ \underbrace{\frac{1}{2}\|Au-b^\delta\|^2}_{\text{data fidelity}} + \underbrace{\lambda\,\mathrm{TV}(u)}_{\text{regularization}}, \tag{6}$$

which is a specific instance of the model problem (1) where $f(u)=\frac{1}{2}\|Au-b^\delta\|^2$, $\lambda>0$ is a regularization parameter, and TV is the *Total Variation* operator defined as

$$\mathrm{TV}(x) = \sum_{i=1}^{d}\|\nabla_i u\|. \tag{7}$$

Methods

Similarly to eqs. (1) and (6), given a positive definite matrix $S\in\mathbb{R}^s$, we consider the optimization problem

$$\underset{u\in\mathbb{R}^d}{\operatorname{argmin}}\ f_S(u) + h(Wu), \tag{8}$$

where

$$f_S(u) = \frac{1}{2}\|S^{-\frac{1}{2}}(Au-b^\delta)\|^2, \tag{9}$$

the positive definite linear operator $S^{\frac{1}{2}}$ can be interpreted as a left preconditioner for the linear system (5).

Assuming that there exists a positive definite matrix $P\in\mathbb{R}^d$ such that

$$P^{-1}A^T = A^T S^{-1}, \tag{10}$$

then it holds

$$\nabla f_S(u) = A^T S^{-1}(Au-b^\delta) = P^{-1}\nabla f(u), \tag{11}$$

where $f(u)=\frac{1}{2}\|Au-b^\delta\|^2$ as in equation (6).

Therefore, applying the NPD algorithm to problem (8), under the assumption (10), we obtain an iterative scheme, named *Preconditioned Nested Primal-Dual* (PNPD). Figure 1, shows the iteration scheme of PNPD and its relationship with the original NPD and NPDIT algorithms.

NPD

$$\begin{cases}\bar{u}_n = u_n + \gamma_n(u_n - u_{n-1}),\\ u_{n+\frac{1}{2}} = \bar{u}_n - \alpha_n\nabla f(\bar{u}_n),\\ u_{n+1} \approx \operatorname{prox}_{\alpha_n h\circ W}(u_{n+\frac{1}{2}}),\end{cases}$$

Left Prec. → PNPD

$$\begin{cases}\bar{u}_n = u_n + \gamma_n(u_n - u_{n-1}),\\ u_{n+\frac{1}{2}} = \bar{u}_n - \alpha_n P^{-1}\nabla f(\bar{u}_n),\\ u_{n+1} \approx \operatorname{prox}_{\alpha_n h\circ W}(u_{n+\frac{1}{2}}),\end{cases}$$

Right Prec. → NPDIT

$$\begin{cases}\bar{u}_n = u_n + \gamma_n(u_n - u_{n-1}),\\ u_{n+\frac{1}{2}} = \bar{u}_n - \alpha_n P^{-1}\nabla f(\bar{u}_n),\\ u_{n+1} \approx \operatorname{prox}^{P}_{\alpha_n h\circ W}(u_{n+\frac{1}{2}}),\end{cases}$$

Figure 1. Iteration scheme of PNPD, NPD, and NPDIT and relationship between the three algorithms.

Preconditioner choice

A reasonable choice for the preconditioner P and the associated matrix S, that satisfies the condition (10), is

$$P = A^TA + \nu I, \qquad S = AA^T + \nu I, \tag{12}$$

with $\nu > 0$. This is inspired by the iterated Tikhonov method, i.e., the Levenberg–Marquardt method applied to linear problems. With this choice, the preconditioner P contains some second-order information about the problem, which can be useful to accelerate the convergence of the algorithm. More in general, the identity in equation (10) is satisfied whenever P is a polynomial of A^TA and S is a corresponding polynomial of AA^T. This particular case is briefly discussed in [1].

Numerical Results

We considered a 256×256 grayscale image of a cameramenin Figure 2a. The blurred image b was obtained using a Gaussian *Point Spread function* (PSF) with $\sigma = 2$ pixels standard deviation. To generate the final observed image b^δ, we added a 1% of white Gaussian noise.

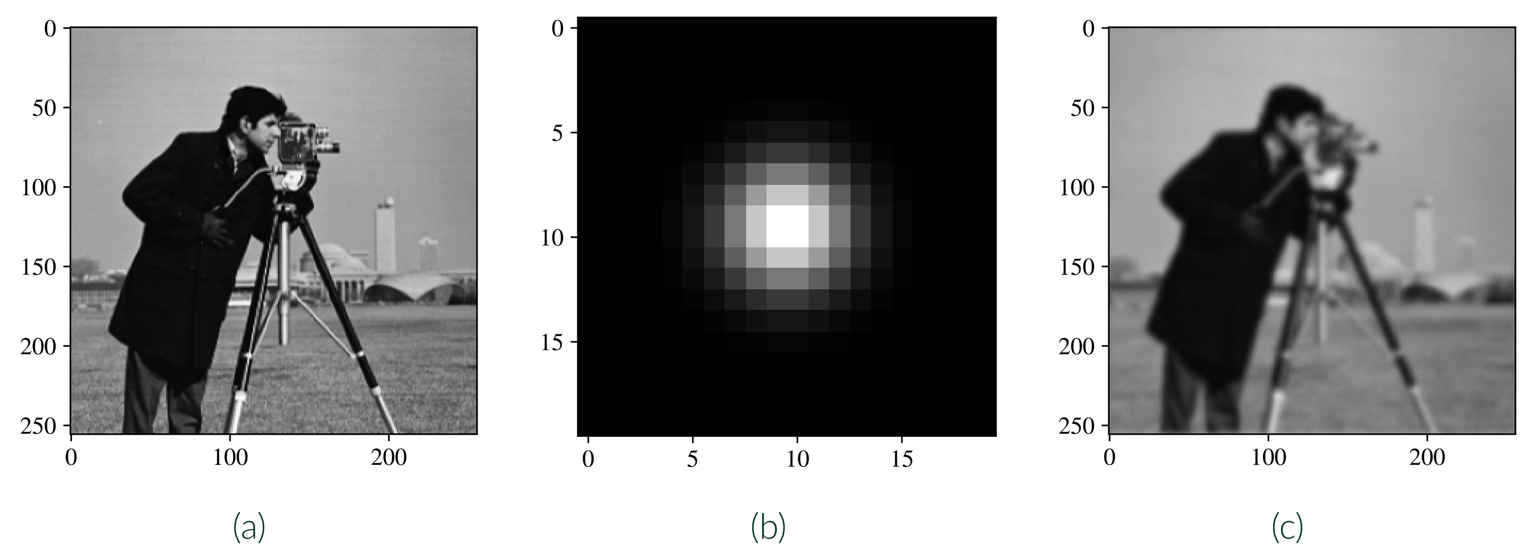

Figure 2. (a) Ground truth image of a cameramen. (b) PSF used to blur the ground truth (center crop of size 20×20). (c) Observed image b^δ obtained by adding white Gaussian noise on top of the discrete circular convolution of the ground truth and the PSF.

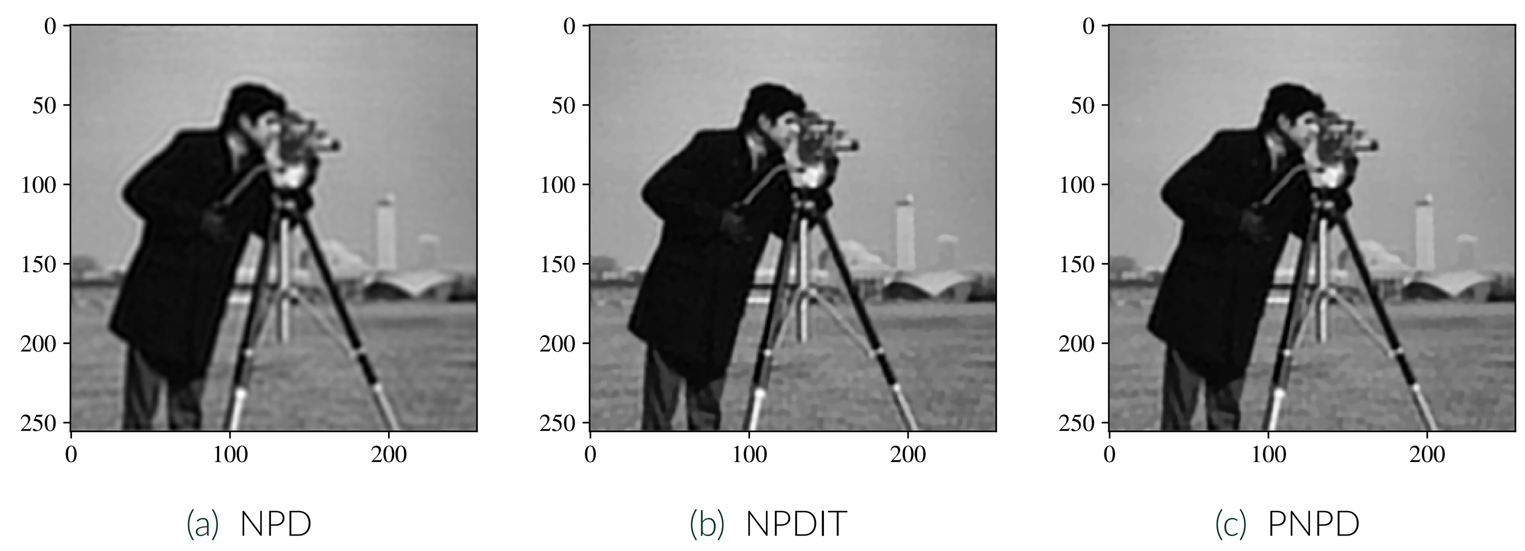

Figure 3. Comparison of the reconstructions obtained with NPD, PNPD, and NPDIT after 10 iterations. The preconditioner parameter is $\nu = 10^{-1}$. The number of nested loop iterations is $k_{\max} = 3$. The regularization parameter is $\lambda = 2\cdot 10^{-4}$ for NPD and NPDIT, while it is $\lambda = 2\cdot 10^{-3}$ for PNPD.

The comparison between the three different algorithms NPD, NPDIT, and PNPD is presented in Figure 4. The performances of each method were measured through the *Structural Similarity Index* (SSIM). Iteration-wise, left plot, we observe that PNPD and NPDIT exhibit similar behaviors, both converging faster than NPD. This is due to the presence of the preconditioner P, which effectively enhances the speed of convergence of the algorithms. In terms of execution time, right plot, the PNPD strategy outperforms both the other methods. The gap between our proposal and NPDIT is due to the fact that as $k_{\max}$ increases, the NPDIT algorithm must compute more FFTs at each iteration. Indeed, the NPDIT method looks for approximate evaluations of $\operatorname{prox}^P_{\alpha h\circ W}$ while PNPD approximates $\operatorname{prox}_{\alpha h\circ W}$ in the same manner as NPD. Therefore, NPDIT has to perform an extra multiplication by P^{-1} for each extra nested iteration compared to PNPD.

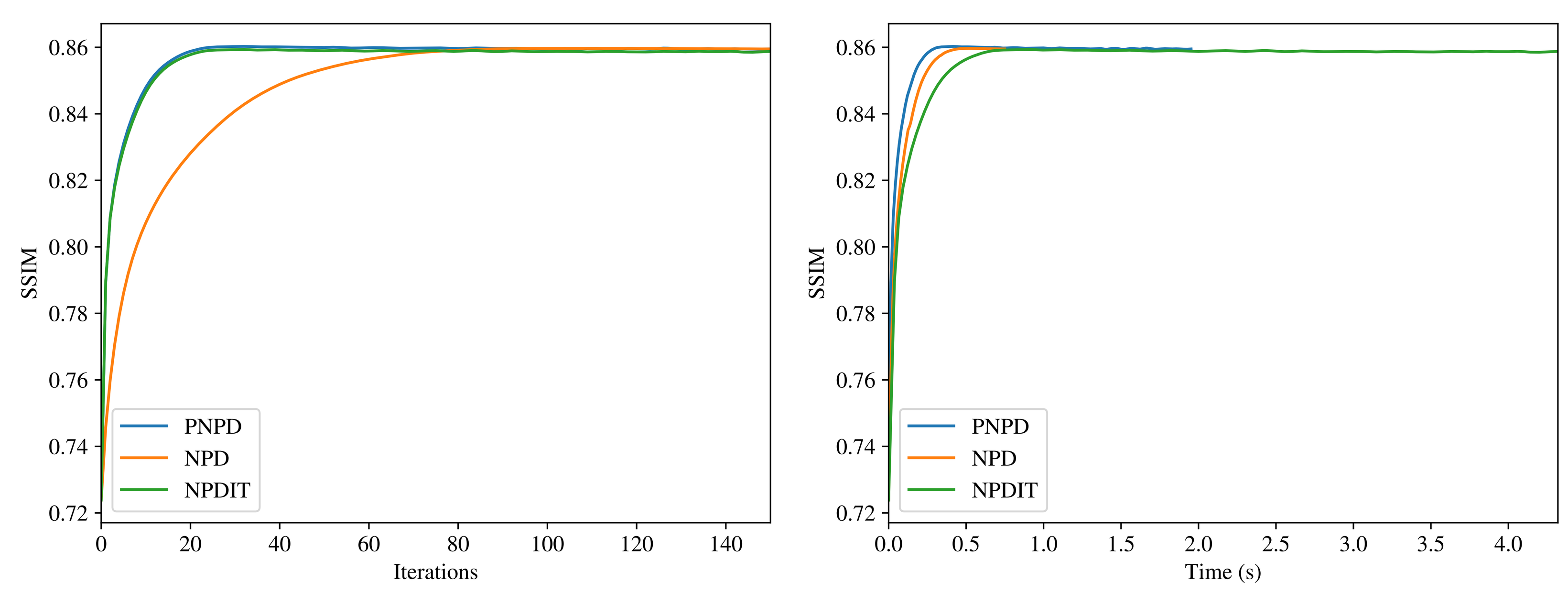

Figure 4. Example 1: Comparison of the SSIMs between PNPD, NPD, and NPDIT. The preconditioner parameter is $\nu = 10^{-1}$. The number of nested loop iterations is $k_{\max} = 1$ for NPD and $k_{\max} = 3$ for NPDIT and PNPD. The regularization parameter is $\lambda = 2\cdot 10^{-4}$ for NPD and NPDIT, while $\lambda = 2\cdot 10^{-3}$ for PNPD.

References

[1] S. Aleotti, M. Donatelli, R. Krause, and G. Scarlato. "A Preconditioned Version of a Nested Primal-Dual Algorithm for Image Deblurring". In: *Journal of Scientific Computing* 103.3 (Apr. 2025), p. 85.

[2] S. Aleotti, S. Bonettini, M. Donatelli, M. Prato, and S. Rebegoldi. "A nested primal–dual iterated Tikhonov method for regularized convex optimization". In: *Computational Optimization and Applications* (Sept. 2024).

[3] S. Bonettini, M. Prato, and S. Rebegoldi. "A nested primal–dual FISTA-like scheme for composite convex optimization problems". In: *Computational Optimization and Applications* 84.1 (Jan. 2023), pp. 85–123.