

The Stabilizing Properties of Nonnegativity Constraints in Image Deblurring Problems

J. Bardsley¹, J.K. Merikoski², and R. Vio³

¹ Department of Mathematical Sciences, The University of Montana, Missoula, MT 59812-0864, USA.
e-mail: bardsleyj@mso.umt.edu

² Department of Mathematics, Statistics and Philosophy, University of Tampere, FI-33014, Finland
e-mail: jorma.merikoski@uta.fi

³ Chip Computers Consulting s.r.l., Viale Don L. Sturzo 82, S.Liberale di Marcon, 30020 Venice, Italy
e-mail: robertovio@tin.it

Received; accepted

ABSTRACT

Aims. It is well known from practice that incorporating nonnegativity constraints in image deblurring algorithms often yields solutions that are much more stable with respect to errors in the data. In the current literature, no formal explanation of the stabilizing effects of nonnegativity constraints has been given. In this paper, we present both theoretical and computational results in support of this empirical finding.

Methods. Our arguments are developed in the context of the least-squares approach. For our analysis, we express the solution of a nonnegatively constrained least squares-problem as a pseudo-solution of a linear system. The conditioning of the corresponding coefficient matrix is then compared with the conditioning of the coefficient matrix of the linear system without constraints.

Results. In general, the matrix corresponding to the nonnegatively constrained problem is better conditioned than that of the unconstrained problem, and as a result, the corresponding solutions are typically more stable with respect to errors in the data.

Conclusions. In astronomical imaging, some form of regularization is either implicitly or explicitly used in all algorithms. The most standard regularization techniques, e.g., Tikhonov and iterative regularization, bias solutions in a way that is not compatible with the true solution. The incorporation of nonnegativity constraints, on the other hand, provides stability in a way that is fully compatible with the true solution. When incorporated into the least-squares framework, the resulting algorithms often yield reconstructions that are either on par, or are of a higher quality, than those obtained with more sophisticated approaches. This suggests that incorporating prior information about the object has a greater impact on results than does the sophistication of the algorithm that is used. We emphasize that this is not an academic result. In fact, simple algorithms such as those based on linear least-squares approaches are easy to implement, more flexible regarding the incorporation of constraints and are available in the most popular packages. Consequently, we believe that in many situations they should be the first choice in deblurring problems.

Key words. Methods: data analysis – Methods: statistical – Techniques: Image processing

1. Introduction

The restoration of images is a common problem in Astronomy. Astronomical images are blurred due to several factors: the refractive effects of the atmosphere, the diffractive effects of the finite aperture of the telescope, the statistical fluctuations inherent in the collection of images by a CCD camera, and instrumental defects. The widespread interest in this subject has resulted in the development of a large number of algorithms with different degrees of sophistication (for a review, see Starck et al, 2002).

In the past, there have been various claims regarding the superiority of a one class of algorithms over an-

other (e.g., nonlinear vs. linear). However, it is necessary to take into account that without the exploitation of a priori information about the true solution, one algorithm cannot be expected to provide better reconstructions than another. In particular, approaches that take into account noise statistics are hindered in image reconstruction by the fact that the underlying problem is ill-conditioned. More specifically, some type of regularization, e.g., Tikhonov or iterative regularization, is typically needed, which amounts to a (more or less explicit) low-pass filtering of the observed images. The filtering operation has an important consequence: it does not permit the full exploitation of the information con-

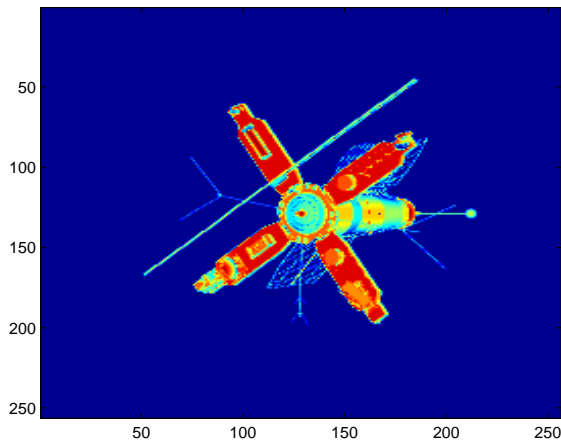


Fig. 1. Satellite object. The image is 256×256 pixels.

tained in the highest Fourier frequencies, i.e., those in which the specific nature of the noise has the largest influence. Consequently, there may be little benefit seen, in terms of reconstruction quality, from the incorporation of the specific statistical noise model, resulting in a leveling in the reconstruction quality obtained by different classes of algorithms. For example, Vio et al. (2005) and Vio & Wamsteker (2005) have shown that the simple, least-squares Landweber (LW) algorithm yields reconstructions of comparable quality to the more sophisticated Richardson-Lucy (RL) algorithm, even in the case of Poissonian noise (recall that RL can be viewed as a maximum-likelihood method in the Poisson noise case). To illustrate this point, in Figs. 1-3 a numerical experiment is presented in which an object is superimposed onto a flat, faint sky-background, and the blurred image is corrupted with Poisson noise. Since RL assumes a more accurate statistical model (NB., here we have used a version of the algorithm that works with a sky-background different from zero - see Bertero & Boccacci 2000), one would expect it to be the superior algorithm. From Fig. 2, it is evident that RL is superior when compared to LW. However, a comparison of RL with the projected (non-negatively constrained after sky-background subtraction) LW algorithm suggests that RL is superior to LW not because it assumes a more accurate statistical model, but because it incorporates the constraint on the level of the sky-background. This supports our belief that in many cases a least-squares approach should be used in place of more sophisticated approaches, since the corresponding algorithms (with or without constraints) are straightforward to implement. For this reason, we will focus our attention on the stabilizing effects of nonnegativity constraints on least-squares problems.

The paper is organized as follows. We will begin in Sect. 2 with a formalization of the problem. In Sect. 3, using basic constrained optimization theory, we express the solution of a nonnegatively constrained least squares problem as a pseudo-solution of a linear system. In Sect. 4,

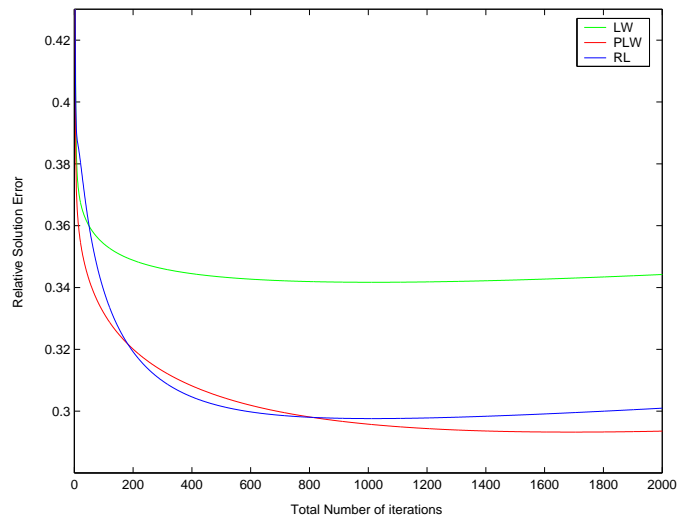


Fig. 2. Plot of the *relative solution error* $\|\mathbf{x}_k - \mathbf{x}_{\text{true}}\|/\|\mathbf{x}_{\text{true}}\|$ for the deblurring obtained with the least-squares *Landweber* (LW), *Projected Landweber* (PLW), and *Richardson-Lucy* (RL) algorithms. The original object, shown in Fig. 1, is superimposed on a background whose level is 1% of the peak value of the blurred image. The PSF is a two-dimensional Gaussian with a dispersion of 8 pixels. The size of the image is 256×256 pixels. The noise is Poissonian with peak S/N = 30 dB.

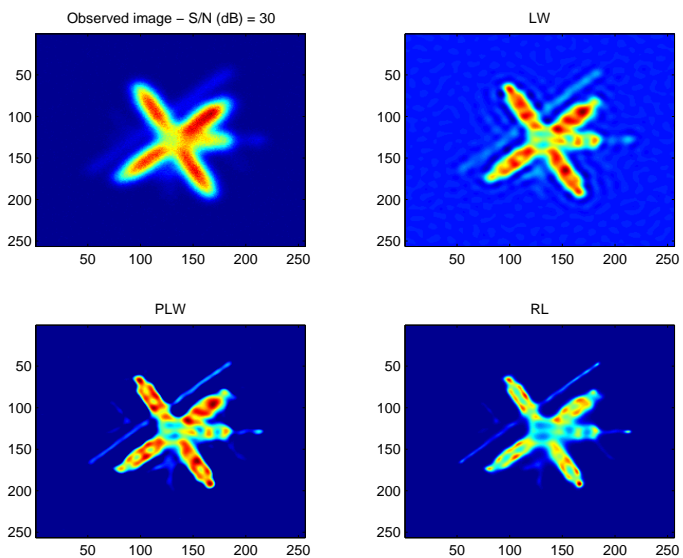


Fig. 3. Solutions corresponding to the minimum of the *relative solution errors* in Fig. 2.

theoretical results regarding the condition numbers of such linear system are presented, and in Sect. 5 our numerical arguments are given. Finally, conclusions are made in Sect. 6.

2. Formalization of the problem

When a two-dimensional object $x(s, t)$ in outer-space is viewed through a telescope the observed image $b(s, t)$ is

related to $x(s, t)$ via the integral equation

$$b(s, t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(s, t, s', t') x(s', t') ds' dt', \quad (1)$$

where $A(s, t, s', t')$ is called the *point-spread function* (PSF) and characterizes the blurring effects of the imaging system of interest. In the case of a *spatially-invariant* PSF, model (1) takes the convolution form

$$b(s, t) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(s - s', t - t') x(s', t') ds' dt'. \quad (2)$$

In many experimental situations, this is the model that is assumed.

Model (2) is only theoretical. In practical applications, we only have discrete observations of the image $b(s, t)$. This means that we have to deal with a discrete linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \quad (3)$$

Here \mathbf{b} and \mathbf{x} are the discrete, column stacked image and object respectively, and \mathbf{A} is the discrete blurring matrix, which is determined by A and is typically ill-conditioned.

We will focus our analysis on $n \times n$ linear systems of the form (3), where \mathbf{A} is nonsingular, and subject to the nonnegativity constraints

$$x_i \geq 0, \quad i = 1, \dots, n. \quad (4)$$

In the sequel, we will use $\mathbf{x} \geq \mathbf{0}$ to denote (4). One approach for estimating the discrete object, or true image, from \mathbf{b} is to solve the least squares minimization problem

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2. \quad (5)$$

Here $\|\cdot\|$ denotes the standard ℓ^2 norm. Since \mathbf{A} is nonsingular, the solution of (5) is given by $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. In this paper, we seek to analyze the effects of replacing (5) by the nonnegatively constrained least squares problem

$$\min_{\mathbf{x} \geq \mathbf{0}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2. \quad (6)$$

Due to the fact that \mathbf{A} is ill-conditioned, solutions of both (5) and (6) are typically unstable with respect to noise in data vector \mathbf{b} . What we will show, via both theoretical arguments and numerical results, is that this instability can be expected to be less pronounced in (6).

Before continuing, it is necessary to stress that with model (6) we are implicitly assuming that the object is superimposed to a black sky-background. In practical applications, this condition is clearly unrealistic. In any case, this point can be trivially solved by using an image $\mathbf{b}^* = \mathbf{b} - \mathbf{s}$, with \mathbf{s} the sky-background (e.g., see Bertero & Boccacci 2000).

3. Nonnegatively Constrained Least Squares

Given our assumption that \mathbf{A} is nonsingular, problem (6) has a unique solution. In this section, using some basic theory from constrained optimization, we express this solution as a pseudo-solution of a linear system.

We begin by making a few definitions. First, we say that \mathbf{x} is *feasible* if $\mathbf{x} \geq \mathbf{0}$, and we define the *active set* of any feasible \mathbf{x} by

$$\mathcal{A}(\mathbf{x}) = \{i \mid x_i = 0\}. \quad (7)$$

We can then define the diagonal matrix $\mathbf{D}(\mathbf{x})$ for $\mathbf{x} \geq \mathbf{0}$ by

$$[\mathbf{D}(\mathbf{x})]_{jj} = \begin{cases} 1, & j \notin \mathcal{A}(\mathbf{x}), \\ 0, & j \in \mathcal{A}(\mathbf{x}). \end{cases} \quad (8)$$

We now focus our attention on (6). Note that

$$\|\mathbf{A}\mathbf{x} - \mathbf{b}\|^2 = \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - 2\mathbf{x}^T \mathbf{A}^T \mathbf{b} + \mathbf{b}^T \mathbf{b}. \quad (9)$$

Since \mathbf{A} is invertible, $\mathbf{A}^T \mathbf{A}$ is positive definite, and hence, the least squares function is strictly convex. Thus, (6) has a unique minimizer. To obtain the form of this minimizer, we note that since (9) is strictly convex, the necessary and sufficient conditions for \mathbf{x}^* to be the unique minimizer of (6) are given by Eq. (4) together with

$$[\mathbf{A}^T \mathbf{A} \mathbf{x}^* - \mathbf{A}^T \mathbf{b}]_i \geq 0, \quad i = 1, \dots, n, \quad (10)$$

$$x_i^* \cdot [\mathbf{A}^T \mathbf{A} \mathbf{x}^* - \mathbf{A}^T \mathbf{b}]_i = 0, \quad i = 1, \dots, n. \quad (11)$$

Equations (4), (10), and (11) are known as the *Karush-Kuhn-Tucker (KKT) conditions* for (6) (see Nocedal & Wright (1999) for details).

From (11), we have that for all i such that $x_i^* > 0$

$$[\mathbf{A}^T \mathbf{A} \mathbf{x}^* - \mathbf{A}^T \mathbf{b}]_i = 0. \quad (12)$$

Then, using the notation $\mathbf{D}^* = \mathbf{D}(\mathbf{x}^*)$, since $\mathbf{D}^* \mathbf{x}^* = \mathbf{x}^*$, we have that

$$\mathbf{D}^* (\mathbf{A}^T \mathbf{A} \mathbf{D}^* \mathbf{x}^* - \mathbf{A}^T \mathbf{b}) = \mathbf{0}, \quad (13)$$

or

$$\mathbf{D}^* \mathbf{A}^T \mathbf{A} \mathbf{D}^* \mathbf{x}^* - \mathbf{D}^* \mathbf{A}^T \mathbf{b} = \mathbf{0}, \quad (14)$$

which are the normal equations for

$$\min_{\mathbf{x}} \|\mathbf{A} \mathbf{D}^* \mathbf{x} - \mathbf{b}\|^2. \quad (15)$$

To find an expression for \mathbf{x}^* , we note that since $x_i^* = 0$ for all $i \in \mathcal{A}(\mathbf{x}^*)$, \mathbf{x}^* is the minimum norm solution of (15). That is,

$$\mathbf{x}^* = (\mathbf{A} \mathbf{D}^*)^\dagger \mathbf{b}, \quad (16)$$

where “ \dagger ” denotes *pseudo-inverse*.

The analysis above shows that the minimum norm solution of the unconstrained problem (15) and the unique solution of the nonnegatively constrained problem (6) are the same. In practice, though, one does not solve (15) since \mathbf{D}^* is not known *a priori*. Nonetheless, a careful consideration of (15) and the associated linear system

$$\mathbf{A} \mathbf{D}^* \mathbf{x} = \mathbf{b}. \quad (17)$$

will be a key component of the arguments set forth in this paper.

4. The Condition Numbers of \mathbf{A} and $\mathbf{A}\mathbf{D}^*$.

We have shown that the imposition of nonnegativity constraints in (6) corresponds to replacing \mathbf{A} in (3) by $\mathbf{A}\mathbf{D}^*$, giving (17). Our objective is to investigate and compare the stability of linear systems (3) and (17). For this we need to define the *singular value decomposition (SVD)* of a matrix.

The SVD of an $m \times n$, $m \geq n$, matrix \mathbf{B} with full column rank (see Björck 1996) is given by

$$\mathbf{B} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \quad \mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_1 \\ 0 \end{pmatrix}, \quad (18)$$

where \mathbf{U} is an $m \times m$ unitary matrix, \mathbf{V} is an $n \times n$ unitary matrix, and $\mathbf{\Sigma}$ is $m \times n$ with $\mathbf{\Sigma}_1 = \text{diag}(s_1, s_2, \dots, s_n)$ and $s_1 \geq s_2 \geq \dots \geq s_n > 0$. The s_i 's are known as the *singular values* of \mathbf{B} .

We now survey some basic facts about singular values (for a thorough exposition, see, e.g., Horn & Johnson 1985, 1991). Let $s_i(\mathbf{B})$ denote the i th largest singular value of \mathbf{B} , and again, suppose \mathbf{B} has full column rank. Then

$$\|\mathbf{B}\| = \max_{\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n} \frac{\|\mathbf{B}\mathbf{x}\|}{\|\mathbf{x}\|} = s_1(\mathbf{B}), \quad (19)$$

and (see Ben-Israel & Greville 1974, Exercise 6.17),

$$\|\mathbf{B}^\dagger\|^{-1} = \min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^n} \frac{\|\mathbf{B}\mathbf{x}\|}{\|\mathbf{x}\|} = s_n(\mathbf{B}). \quad (20)$$

The *condition number* of \mathbf{B} is then defined

$$\text{cnd } \mathbf{B} = s_1(\mathbf{B})/s_n(\mathbf{B}) = \|\mathbf{B}\|\|\mathbf{B}^\dagger\|. \quad (21)$$

Although a direct relationship between $\text{cnd } \mathbf{B}$ and the instability of solutions of linear systems of the form $\mathbf{B}\mathbf{x} = \mathbf{d}$ to small changes in \mathbf{d} does not exist, it is typically the case that the larger $\text{cnd } \mathbf{B}$ is, the greater is this instability (see Hansen 1997). This motivates the approach that we will now take in our arguments. Namely, we will show that nonnegativity constraints have a stabilizing effect by showing that, in general, we can expect $\mathbf{A}\mathbf{D}^*$ to have a smaller condition number than \mathbf{A} .

First, note that there exists a permutation matrix \mathbf{Q} such that the first r diagonal entries of the diagonal matrix $\mathbf{D}' = \mathbf{Q}^T\mathbf{D}^*\mathbf{Q}$ are one, with the remaining diagonal entries zero. We define $\mathbf{A}' = \mathbf{A}\mathbf{Q}$. Then the singular values of $\mathbf{A}'\mathbf{D}'$ are equal to the singular values of $\mathbf{A}\mathbf{D}^*$. Therefore, without loss of generality, we may assume that

$$\mathbf{D}^* = \begin{pmatrix} \mathbf{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \quad (22)$$

Consider the partition $\mathbf{A} = (\mathbf{A}_1 \mathbf{A}_2)$, where the first part consists of the first r columns and the second part of the remaining columns of \mathbf{A} . Then $\mathbf{A}\mathbf{D}^* = (\mathbf{A}_1 \mathbf{0})$. Since \mathbf{A} has full column rank, so does \mathbf{A}_1 . Thus $\text{cnd } \mathbf{A}_1$ is defined. We will show that $\text{cnd } \mathbf{A}_1 \leq \text{cnd } \mathbf{A}$. First note that $r \leq n$ (recall that \mathbf{A} is $n \times n$). Then if $\mathbf{x} \in \mathbb{R}^r$, and we denote

$$\tilde{\mathbf{x}} = \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix}, \quad (23)$$

we have

$$\begin{aligned} s_1(\mathbf{A}_1) &= \max_{\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^r} \frac{\|\mathbf{A}_1\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{\mathbf{0} \neq \tilde{\mathbf{x}} \in \mathbb{R}^n} \frac{\|\mathbf{A}\tilde{\mathbf{x}}\|}{\|\tilde{\mathbf{x}}\|} \\ &\leq \max_{\mathbf{0} \neq \mathbf{z} \in \mathbb{R}^n} \frac{\|\mathbf{A}\mathbf{z}\|}{\|\mathbf{z}\|} = s_1(\mathbf{A}), \end{aligned} \quad (24)$$

and

$$\begin{aligned} s_r(\mathbf{A}_1) &= \min_{\mathbf{0} \neq \mathbf{x} \in \mathbb{R}^r} \frac{\|\mathbf{A}_1\mathbf{x}\|}{\|\mathbf{x}\|} = \min_{\mathbf{0} \neq \tilde{\mathbf{x}} \in \mathbb{R}^n} \frac{\|\mathbf{A}\tilde{\mathbf{x}}\|}{\|\tilde{\mathbf{x}}\|} \\ &\geq \min_{\mathbf{0} \neq \mathbf{z} \in \mathbb{R}^n} \frac{\|\mathbf{A}\mathbf{z}\|}{\|\mathbf{z}\|} = s_n(\mathbf{A}). \end{aligned} \quad (25)$$

Therefore

$$\text{cnd } \mathbf{A}_1 = \frac{s_1(\mathbf{A}_1)}{s_r(\mathbf{A}_1)} \leq \frac{s_1(\mathbf{A})}{s_n(\mathbf{A})} = \text{cnd } \mathbf{A}. \quad (26)$$

If we define the condition number of $\mathbf{A}\mathbf{D}^*$ by

$$\text{cnd } \mathbf{A}\mathbf{D}^* := \text{cnd } \mathbf{A}_1, \quad (27)$$

inequality (26) yields

$$\text{cnd } \mathbf{A}\mathbf{D}^* \leq \text{cnd } \mathbf{A}. \quad (28)$$

Thus solutions of (17) can be expected to be more stable with respect to noise in the data vector \mathbf{b} than are solutions of (3).

In astronomical imaging examples, though, a stronger inequality often holds. Due to the fact that there is often a substantial black or faint background in the astronomical objects being viewed, r is often much smaller than n . In such cases, the vectors $\tilde{\mathbf{x}}$ will cover a small subspace of \mathbb{R}^n . Since the maximum (respectively minimum) of a function over a big set is usually much larger (respectively smaller) than that over a small subset, it is not unusual for $s_1(\mathbf{A}_1) \ll s_1(\mathbf{A})$ in (24) and/or $s_r(\mathbf{A}_1) \gg s_n(\mathbf{A})$ in (25), which yields $s_1(\mathbf{A}_1)/s_r(\mathbf{A}_1) \ll s_1(\mathbf{A})/s_n(\mathbf{A})$. In such cases, (28) can be replaced by

$$\text{cnd } \mathbf{A}\mathbf{D}^* \ll \text{cnd } \mathbf{A}. \quad (29)$$

When (29) holds, one can expect that the solutions of (17), and hence (6), will be much more stable with respect to error in \mathbf{b} than (3). The fact that (29) does in fact hold in practical examples is confirmed by our numerical results in Sect. 5.

5. Numerical Results

In this section some numerical experiments are presented that support our claims. Although the theoretical arguments presented above do not depend on the dimensionality of the problem, here we consider only the one-dimensional version of model (2). This is because only in this instance are we able to compute the condition number of \mathbf{A} directly. Moreover, we have deliberately chosen non-astronomical objects. In fact, because of the more important contribution of the high frequency components, objects with sharp outlines are much more difficult to restore

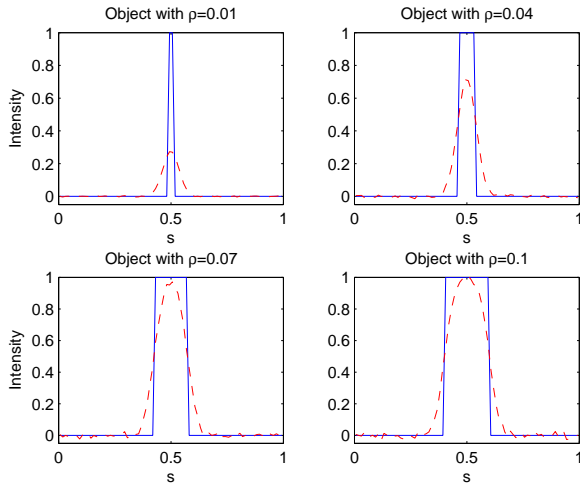


Fig. 4. Plots of Objects and Blurred, Noisy Images. The plot of the discrete true image \mathbf{x}_{true} has the solid line and is defined by (34). The plot of the true image corresponding to $\rho = 0.01$ is given in the upper left, to $\rho = 0.04$ in the upper right, to $\rho = 0.07$ in the lower left, and to $\rho = 0.1$ in the lower right. The plots of the corresponding blurred, noisy images have the dashed line.

than those of typical astronomical observations. Hence, we consider very sharply structured objects.

The one-dimensional version of model (2) is given by

$$b(s) = \int_0^1 A(s - s')x(s')ds', \quad (30)$$

where we use a PSF $A(s)$ given by a Gaussian

$$A(s) = C \exp(-s^2/2\gamma^2), \quad (31)$$

with C and γ positive parameters.

Discretizing (30) yields a linear system of the form (3). If midpoint quadrature is used in the s' variable, the resulting matrix \mathbf{A} has the form

$$[\mathbf{A}]_{ij} = hC \exp\left(-\frac{(i-j)^2 h^2}{2\gamma^2}\right), \quad 1 \leq i, j \leq n, \quad (32)$$

where $h = 1/n$. With this definition, \mathbf{A} is an invertible matrix, and hence, the analysis from the previous sections applies.

We now consider some specific examples. We build our matrix \mathbf{A} using (32) with $n = 80$ and $\gamma^2 = 0.00125$. The data vector \mathbf{b} is generated \mathbf{b} via the statistical model

$$\mathbf{b} = \mathbf{A}\mathbf{x}_{\text{true}} + \mathbf{N}, \quad (33)$$

where \mathbf{x}_{true} is the *object*, or *true image*, and \mathbf{N} is an $n \times 1$ independent and identically distributed, zero mean Gaussian random vector with standard deviation chosen so that the signal-to-noise ratio ($\|\mathbf{b}\|/\|\mathbf{N}\|$) is 30.

We consider several different choices for \mathbf{x}_{true} since \mathbf{D}^* , and hence, the condition number of \mathbf{AD}^* , depends upon the object \mathbf{x}_{true} . We begin by considering objects that have the form

$$\mathbf{x}_{\text{true}} = \begin{cases} 1 - \frac{1}{2} - \rho \leq [\mathbf{x}_{\text{true}}]_i \leq \frac{1}{2} + \rho, \\ 0 \text{ otherwise.} \end{cases} \quad (34)$$

Table 1. Condition Numbers of \mathbf{AD}^* for Various ρ

ρ	$\text{cnd } \mathbf{AD}^*$	s_1	s_r
0.01	2.448×10^2	3.821×10^{-5}	1.561×10^{-7}
0.04	6.589×10^4	5.784×10^{-5}	8.780×10^{-10}
0.07	2.330×10^4	5.392×10^{-5}	2.313×10^{-10}
0.1	1.678×10^4	5.115×10^{-5}	3.048×10^{-10}

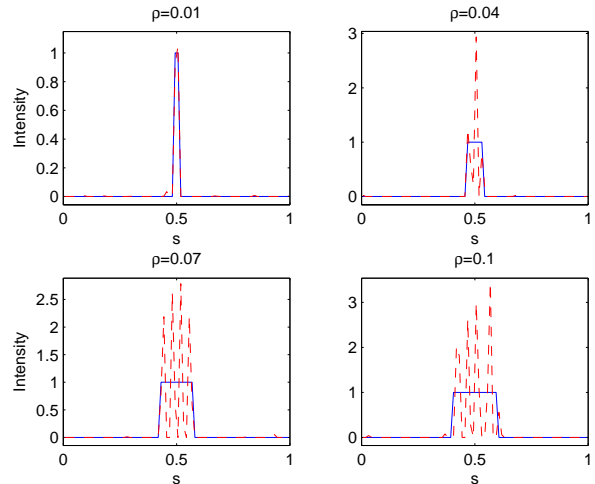


Fig. 5. Plots of the Minimizers of $\|\mathbf{Ax} - \mathbf{b}\|^2$ Subject to the Constraint $\mathbf{x} \geq \mathbf{0}$. The plot of the discrete true image \mathbf{x}_{true} has the solid line. The nonnegatively constrained solutions have the dashed line. The upper left-hand corner corresponds to the true image with $\rho = 0.01$, the upper right to the true image with $\rho = 0.04$, the lower left to the true image with $\rho = 0.07$, and the lower right to the true image with $\rho = 0.1$.

where $0 \leq \rho \leq \frac{1}{2}$. We perform experiments for four separate values of ρ : $\rho = 0.01$, $\rho = 0.04$, $\rho = 0.07$, and $\rho = 0.1$. The plots of these objects and of the corresponding blurred, noisy images are given in Fig. 4.

An accurate estimate of \mathbf{D}^* for each example can be obtained using the projected Newton (PN) method (Bersekas 1982; Kelley 1999). This method is guaranteed to converge to the unique minimizer of (6) for any initial guess. It can also be shown that there exists an integer N such that $\mathbf{D}(\mathbf{x}_k^{\text{PN}}) = \mathbf{D}^*$ for all $k \geq N$, where \mathbf{x}_k^{PN} is the k th PN iterate. Motivated by this, in each example we take $\mathbf{D}^* = \mathbf{D}(\mathbf{x}_{200}^{\text{PN}})$. From this, $\text{cnd } \mathbf{AD}^* = s_1/s_r$ can be computed. The values of $\text{cnd } \mathbf{AD}^*$, s_1 and s_r for the different objects can be found in Table 1. Note that in all cases, the condition number of \mathbf{AD}^* is substantially less than the condition number of \mathbf{A} , which is 1.2×10^{17} . Thus inequality (29) holds in these cases. Hence, more stable results for the constrained problem can be expected. This is confirmed by Figs. 5 and 6, which show plots of the solutions of the unconstrained least squares problem (5) and of the nonnegatively constrained least squares problem (6) respectively. From these figures it is evident that for these examples the nonnegatively constrained solutions are far more stable with respect to errors in the data than are the unconstrained solutions. We emphasize that the plots

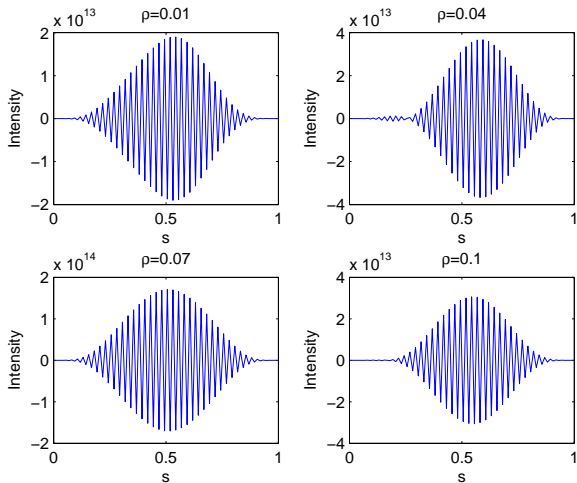


Fig. 6. Plots of the Minimizers of $\|\mathbf{Ax} - \mathbf{b}\|^2$. The upper left-hand corner corresponds to the true image with $\rho = 0.01$, the upper right to the true image with $\rho = 0.04$, the lower left to the true image with $\rho = 0.07$, and the lower right to the true image with $\rho = 0.1$.

in Fig. 5 should not be viewed as reconstructions. They are presented only to illustrate the benefit, in terms of stability, of imposing the non-negativity constraint on the solution.

We next consider an example in which the object \mathbf{x}_{true} consists of several sources. A plot of both \mathbf{x}_{true} and \mathbf{b} is given in Fig. 7. In this case, the object is composed of three sources, each of the type (34): one with $\rho = 0$, one with $\rho = 0.01$ and one with $\rho = 0.1$. For this example, with \mathbf{D}^* computed as in the previous examples, $\text{cnd } \mathbf{AD}^* \approx 1.069 \times 10^5$. Since $\text{cnd } \mathbf{A} \approx 1.2 \times 10^{17}$, inequality (29) is satisfied for this example as well. Hence, more stable results for the nonnegatively constrained problem can be expected. A plot of the nonnegatively constrained least-squares solution is given in Fig. 8. Again, notice that noise amplification occurs, but a comparison with the unconstrained least-squares solution, which is analogous to those found in Fig. 6, shows that the amplification is mild.

6. Conclusions

The deblurring problems that arise in image reconstruction are inherently ill-conditioned. Exact solutions of such problems are typically unusable due to the fact that they are unstable with respect to errors in the collected image. For this reason, some form of regularization of the solution is necessary. The most standard regularization techniques, e.g., Tikhonov and iterative regularization, bias solutions in a way that is not compatible with the true solution. We have provided both theoretical and numerical arguments to the effect that the imposition of nonnegativity constraints can be expected, in many cases, to improve stability, but in a way that is fully compatible with the true solution. We focused our attention on the stabilizing

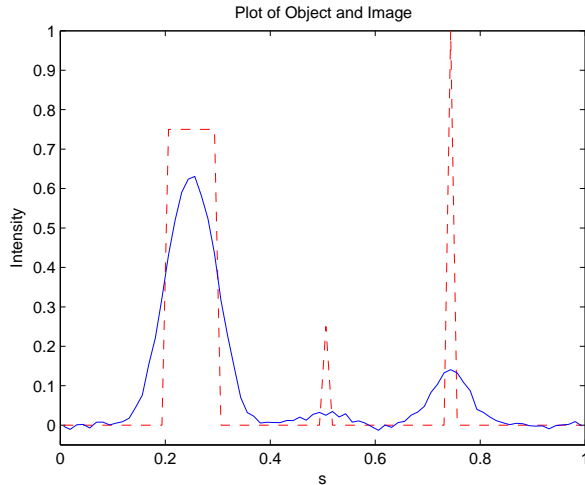


Fig. 7. Plots of an Object and Corresponding Blurred, Noisy Image. The plot of the discrete object \mathbf{x}_{true} has the dashed line. The plot of the noisy, blurred image \mathbf{b} has the solid line.

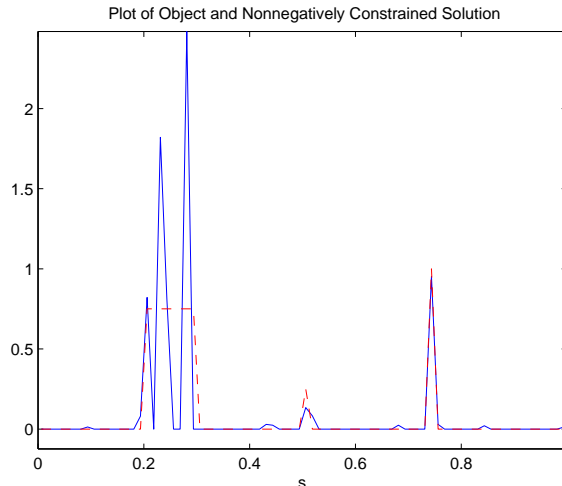


Fig. 8. Plots of the True Image and the Minimizer of $\|\mathbf{Ax} - \mathbf{b}\|^2$ Subject to the Constraint $\mathbf{x} \geq \mathbf{0}$. The plot of the discrete true image \mathbf{x}_{true} has the dashed line. The plot of the nonnegatively constrained solution has the solid line.

effects of nonnegativity constraints in linear least-squares problems. This choice allowed for a more straightforward analysis. However, the fact that when the nonnegativity constraint on the solution is not active, the (nonlinear) Richardson-Lucy algorithm and the (linear least-squares) Landweber algorithm provide quite similar results (Vio et al. 2005; Vio & Wamsteker 2005), leads us to conjecture that nonnegativity constraints will also provide stability for other approaches and that, in general, approaches that are more sophisticated than least-squares cannot be expected to yield better results. More generally speaking, these results suggest that the incorporation of a priori information is more important than the sophistication of the approach that is taken, and hence, that the simple least-squares approach should be the first choice for this kind

References

- Ben-Israel, A., & Greville, T.N.E. 1974, *Generalized Inverses: Theory and Applications* (New York: Wiley)
- Bertsekas, D.P. 1982, *SIAM Journal on Control and Optimization*, 20, 221
- Bertero, M., & Boccacci, P. 2000, *A&A*, 147, 323
- Björck, A. 1996, *Numerical Methods for Least Squares Problems* (Philadelphia: SIAM)
- Hansen, P.C. 1997, *Rank-Deficient and Discrete Ill-Posed Problems* (Philadelphia: SIAM)
- Horn, R.A., & Johnson, C.R. 1985, *Matrix Analysis* (Cambridge: Cambridge University Press)
- Horn, R.A., & Johnson, C.R. 1991, *Topics in Matrix Analysis* (Cambridge: Cambridge University Press)
- Kelley, C.T., 1999, *Iterative Methods for Optimization* (Philadelphia: SIAM).
- Nocedal, J., & Wright, S.J. 1999, *Numerical Optimization* (New York: Springer)
- Starck, J.L., Pantin, E. & Murthag, F. 2002, *PASP*, 114, 1051
- Vio, R., Bardsley, J., & Wamsteker, W. 2005, *A&A*, 436, 741
- Vio, & Wamsteker, W. 2005, *A&A*, 439, 1229