Constrained variable projection method for blind deconvolution

A Cornelio\textsuperscript{1}, E Loli Piccolomini\textsuperscript{2}, J G Nagy\textsuperscript{3}

\textsuperscript{1}Department of Pure and Applied Mathematics, University of Modena and Reggio Emilia
\textsuperscript{2}Department of Mathematics, University of Bologna
\textsuperscript{3}Department of Mathematics and Computer Science, Emory University

E-mail: anastasia.cornelio@unimore.it

Abstract. This paper is focused on the solution of the blind deconvolution problem, here modeled as a separable nonlinear least squares problem. The well known ill-posedness, both on recovering the blurring operator and the true image, makes the problem really difficult to handle. We show that, by imposing appropriate constraints on the variables and with well chosen regularization parameters, it is possible to obtain an objective function that is fairly well behaved. Hence, the resulting nonlinear minimization problem can be effectively solved by classical methods, such as the Gauss-Newton algorithm.

1. Introduction
Blind deconvolution can be modeled as an inverse problem of the form \[ b = A(y_{\text{true}})x_{\text{true}} + \eta \] (1)
where \( b \in \mathbb{R}^m \) is the measured, blurred and noisy image (\( \eta \) models the noise), and \( x_{\text{true}} \in \mathbb{R}^n \) represents the unknown true image. The vector \( y_{\text{true}} \in \mathbb{R}^p \) is unknown and \( A(y) \) is a nonlinear operator, which models the blurring and maps \( y \) into an \( m \times n \) matrix. We call equation (1) a separable inverse problem because the measured data depends linearly on the unknown vector \( x \) and nonlinearly on the unknown vector \( y \).

In [1] the authors, once defined the model mapping \( A(\cdot) \) and given the noisy data \( b \), formulate the optimization problem
\[
\min_{x, y} f_0(x, y) 
\] (2)
and then propose to apply the variable projection method [2, 3, 4, 5] in order to determine the unknowns \( x \) and \( y \). By taking advantage of the separability of the variables, the method consists on eliminating the linear variables \( x \) thus obtaining a reduced nonlinear cost functional, that depends only on \( y \)
\[
f(y) = f_0(x(y), y). 
\] (3)
Aim of this work is to show that imposing appropriate constraints on the unknowns provides a better behaved objective function, which results in computing a better solution of problem (2) when using a standard optimization method such as Gauss-Newton. In particular, we imposed
nonnegative constraints on the linear variables $x$ and stated the problem as a least squares problems with Tikhonov regularization on $x$, that is
\[
\min_{x,y} f_0(x, y) = \| A(y)x - b \|_2^2 + \lambda \| x \|_2^2 \quad \text{subject to } x \geq 0.
\] (4)

Equation (2) is still well defined if linear constraints are imposed on $x$. In [5] Sima and Van Huffel exploit this property and provide results for an application in spectroscopy. Since the problem is well conditioned in this framework, regularization is not needed. The paper is organized as follows. In Section 2 the variable projection method is extended to the nonnegatively constrained case; in Section 3 we investigate the behaviour of the objective function in the presence of constraints, while in Section 4 we provide some experiments. Finally, conclusions are given in Section 5.

2. Variable Projection for nonnegative constraints

The variable projection method [2, 3, 5] has been applied to blind deconvolution modeled as a separable nonlinear least squares problem in [1]. In the following, we will extend the method to the case of nonnegative constraints imposed on $x$.

If we knew the nonlinear parameters $y$, then the linear variables $x(y)$ could be obtained by solving the nonnegative linear least-squares problem:
\[
x(y) = \min_{x \geq 0} \| A(y)x - b \|_2^2 + \lambda \| x \|_2^2.
\] (5)

By replacing (5) in the original objective function (4), we obtain a new reduced cost function, depending only on $y$
\[
\min_y f(y) = f_0(x(y), y).
\] (6)

Now, like in any nonnegatively constrained optimization problem, for any $x \geq 0$ we can define: the active set by $A(x) = \{ i | x_i = 0 \}$ and a diagonal matrix $D(x)$ by
\[
D(x)_{ii} = \begin{cases} 1 & \text{if } i \notin A(x) \\ 0 & \text{if } i \in A(x) \end{cases}.
\]

In the following we will use the notation $D = D(x(y))$, where $x(y)$ is defined by (5). Therefore the solution of (5) can be written by a closed formula as in [6]:
\[
x(y) = \begin{bmatrix} A(y)D \\ \lambda I_n \end{bmatrix}^\dagger \begin{bmatrix} b \\ 0 \end{bmatrix}
\] (7)

where $^\dagger$ indicates the pseudoinverse of a matrix and $0$ is the $n$-dimensional zero column vector. By using (7), (6) becomes exactly
\[
\min_y f(y) = \| A(y) \begin{bmatrix} A(y)D \\ \lambda I_n \end{bmatrix}^\dagger \begin{bmatrix} b \\ 0 \end{bmatrix} - b \|_2^2
\] (8)

which can be solved with a classical nonlinear least squares method, such as the Gauss-Newton algorithm. The computation of the Jacobian $J$ of $f(y)$ is widely discussed in [7].
3. Role of the constraints

In this section we study the behavior of the objective function $f(y)$ defined in (8), in order to investigate the effects of the constraints on the solution of the outer nonlinear problem. To simplify the representation of the function, we consider $y$ as containing just one parameter. In particular, in the following tests we consider the 2-dimensional Gaussian PSF defined as

$$p_{ij} = \exp\left(-\frac{1}{2} \left[ \begin{array}{c} i-k \\ j-l \end{array} \right]^T \left[ \begin{array}{cc} \sigma & 0 \\ 0 & \sigma \end{array} \right]^{-1} \left[ \begin{array}{c} i-k \\ j-l \end{array} \right] \right)$$

(9)

where $y = (\sigma, \sigma)$, $\sigma$ determines the spread of the PSF in both the directions and $(k, l)$ is the center of the PSF.

The goal of this test is to show that the nonnegative constraints imposed on $x$ have a crucial role in our application. In the following, we use the satellite image (Figure 2 (b)), blurred by the Gaussian PSF (9) with $\sigma = 2.5$ and corrupted by 5% Gaussian white noise.

We compare the function $f(y)$ in (8) with the function $g(y)$ obtained without imposing constraints on $x$, i.e.:

$$g(y) = \left\| A(y) \left[ \begin{array}{c} A(y) \\ \lambda I_n \end{array} \right]^{\dagger} \left[ \begin{array}{c} b \\ 0 \end{array} \right] - b \right\|^2_2,$$

(10)

In Figure 1 we compare the objective functions $g(y)$ and $f(y)$ in this easier 1-dimensional case. Let us start with the unconstrained formulation of the problem (Figure 1 (a)). We solve the inner linear problem in (10) with the unconstrained hybrid (HyBR) method proposed in [1]. By using the regularization parameter $\lambda$ automatically chosen by HyBR, we obtain that $g(y)$ is a monotonically increasing function (black dashdot line). Then we use the optimal $\lambda$, obtained when the true $x$ solution is given: now the function is convex (red plain line), but has an unique global minimum that does not match the true value $y = 2.5$.

Let us consider now the nonnegative function $f(y)$ (Figure 1 (b)), where the inner linear problem has been solved by a scaled gradient projection (SGP) method [8]. Since we are not provided by a rule to choose the regularization parameter in the presence of nonnegative constraints, we use the optimal $\lambda$ obtained by the previous unconstrained method (red plain line). Let us remind that imposing nonnegative constraints actually reduces the ill-conditioning of the problem [6]. For this reason it makes sense using, in this case, a smaller $\lambda$ than in the unconstrained problem: the global minimum matches the true solution (blue dashed line).
4. Numerical results
In this section we show how imposing nonnegative constraints on $x$ provides better results for both the linear and nonlinear variables. In particular we consider as true objects the vectorized $256 \times 256$ galaxy and the satellite images (Figure 2), blurred by the Moffat PSF

$$p_{ij} = \left( 1 + \frac{(i - k)^T \sigma_1\sigma_0^{-1}(i - k)}{(j - l)^T \sigma_1\sigma_0^{-1}(j - l)} \right)^{-\beta}$$

defined by the parameter $y_{\text{true}} = (\sigma_1, \sigma_2) = (1,2)$ and corrupted by different levels of white Gaussian noise. The regularization parameter $\lambda$ has been fixed for all the Gauss-Newton iterations and it has been heuristically chosen as the parameter providing the best results in terms of relative error. In Table 1 we compare the unconstrained (10) and the constrained (8) formulation of the blind deconvolution problem. As indicators of the reconstruction quality, the minimum of the $x$ and $y$ relative errors and the structural similarity index [9] (SSIM) are shown.

![True objects](image)

**Figure 2.** True objects

| Problem         | $\text{err}_x$ 3% noise | $\text{err}_y$ 3% noise | SSIM 3% noise | $\text{err}_x$ 5% noise | $\text{err}_y$ 5% noise | SSIM 5% noise |
|-----------------|--------------------------|--------------------------|---------------|--------------------------|--------------------------|---------------|
| Unconstrained: min $g$ | 0.197159                 | 0.225329                 | 0.44505       | 0.218703                 | 0.196787                 | 0.37440       |
| Constrained: min $f$   | 0.072446                 | 0.031953                 | 0.83498       | 0.076305                 | 0.033387                 | 0.81646       |

Moreover, in Figure 3 we consider the satellite image corrupted by 3% noise and blurred with $y = (1.5, 2)$; we compare the relative errors obtained by solving the unconstrained problem (10) (black triangles) and of the constrained problem (8) (magenta squares). Figure 3 (c) confirms that the function $g(y)$ is still decreasing, but the minimum we’re going to reach is far from the ausplicable solution of the problem. On the other hand, $f(y)$ becomes flat during the iterations: this suggests a stopping criterion for the Gauss-Newton algorithm, that involves the objective function, such as

$$\frac{|f(x_{k+1}) - f(x_k)|}{f(x_{k+1})} \leq \frac{\eta}{2}$$

We used as a good value for $\eta$ the level of noise in the observed data.
5. Conclusions

In this paper we have illustrated that a substantial amount of uncertainty arises when solving blind deconvolution problems, firstly caused not only by noise and measurement errors, but also by imprecise regularization parameters and linear solvers. We showed that imposing constraints on some parameters, i.e. solving the ill-posed problem with greater accuracy, produces an objective function more suitable for the common optimization algorithms, such as Gauss-Newton. This suggests that a priori knowledge about the application is useful to include appropriate constraints on $x$ and $y$. However, to efficiently solve this kind of separable ill-posed inverse problems, a substantial amount of work is needed in order to have a robust method. First of all, the development of methods to automatically choose regularization parameters for constrained deconvolution problems could improve significantly the solution of this kind of nonlinear inverse problems.

References
[1] Chung J and Nagy J 2010 SIAM J. Sci. Comput. 31 4654–4674
[2] Golub G H and Pereyra V 1973 SIAM J. Numer. Anal. 10 413–432
[3] Kaufman L 1975 BIT 15 49–57
[4] Kaufman L and Pereyra V 1978 SIAM J. Numer. Anal. 15 12–20
[5] Sima D M and Huffel S V 2007 J. Comput. Appl. Math. 203 264–278
[6] Bardsley J M, Merikoski J and Vio R 2008 IJPAM 43 95–109
[7] Cornelio A, Piccolomini E L and Nagy J G 2012 Numerical optimization methods for blind deconvolution Tech. rep. URL http://www.optimization-online.org/DB_HTML/2012/02/3351.html
[8] Bonettini S, Zanella R and Zanni L 2009 Inverse Problems 25
[9] Wang Z, Bovik A, Sheikh H and Simoncelli E 2004 IEEE Trans. Image Proc. 13 600–612