A generalized global Arnoldi method for ill-posed matrix equations

A. Bouhamidi[a], K. Jbilou[a], L. Reichel[b], H. Sadok[a]

[a] *Université de Lille Nord de France, L.M.P.A, ULCO, 50 rue F. Buisson BP699, F-62228 Calais-Cedex, France*
[b] *Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA*

Abstract

This paper discusses the solution of large-scale linear discrete ill-posed problems with a noise-contaminated right-hand side. Tikhonov regularization is used to reduce the influence of the noise on the computed approximate solution. We consider problems in which the coefficient matrix is the sum of Kronecker products of matrices and present a generalized global Arnoldi method, that respects the structure of the equation, for the solution of the regularized problem. Theoretical properties of the method are shown and applications to image deblurring are described.

Keywords: global Arnoldi, Kronecker product, matrix equation, regularization.

---

1. Introduction

We are concerned with the solution of linear discrete ill-posed matrix equations of the form

$$\sum_{i=1}^{p} A_i \, X \, B_i = G, \tag{1}$$

where the matrices $A_i, B_i \in \mathbb{R}^{n\times n}$ are of ill-determined rank; in particular, they are severely ill-conditioned and may be singular. If the matrix equation

Email addresses: `bouhamidi@lmpa.univ-littoral.fr` (A. Bouhamidi), `jbilou@lmpa.univ-littoral.fr` (K. Jbilou), `reichel@math.kent.edu` (L. Reichel), `sadok@lmpa.univ-littoral.fr` (H. Sadok)

is inconsistent, then (1) should be considered a minimization problem in which the Frobenius norm $\|\cdot\|_F$ of the residual error is minimized.

Let the vector $x = \mathrm{vec}(X) \in \mathbb{R}^{n^2}$ be obtained by stacking the columns of the matrix X. Equation (1) can be expressed as a linear system of equations for the vector x with a matrix that is a sum of Kronecker products of the matrices B_i^T and A_i,

$$\left(\sum_{i=1}^{p} B_i^T \otimes A_i\right) x = \mathrm{vec}(G), \tag{2}$$

where the superscript ^{T} denotes transposition and the Kronecker product of two matrices $C \in \mathbb{R}^{m_c \times n_c}$ and $D \in \mathbb{R}^{m_d \times n_d}$ is defined as the $(m_c m_d) \times (n_c n_d)$ matrix $C \otimes D = [c_{ij} D]$; see, e.g., [10]. The representation (2) can be helpful for determining properties of equation (1); however, we are interested in developing a solution method that respects the structure of equation (1).

The right-hand side matrix $G \in \mathbb{R}^{n \times n}$ in (1) represents available data, such as a blurred and error-contaminated image. The error E in G is assumed to be additive and unknown. We will refer to E as "noise." Thus,

$$G = \widehat{G} + E, \tag{3}$$

where $\widehat{G}$ represents an unknown error-free matrix associated with G.

Let $\widehat{X}$ denote the solution of minimal Frobenius norm of the unavailable matrix equation with the error-free right-hand side,

$$\sum_{i=1}^{p} A_i \widehat{X} B_i = \widehat{G},$$

which we for notational simplicity assume to be consistent. We would like to determine an approximation of $\widehat{X}$ by computing a suitable approximate solution of the available equation with contaminated right-hand side (1). However, straightforward solution of (1) typically does not yield a meaningful approximation of $\widehat{X}$ due to the error E in G and the ill-conditioning of the matrices A_i and B_i. Tikhonov regularization remedies this difficulty by replacing the matrix equation (1) by a minimization problem, whose solution is less sensitive to the error E than the solution of (1). We consider the minimization problem

$$\min_{X \in \mathbb{R}^{n \times n}} \left\{ \left\| \sum_{i=1}^{p} A_i X B_i - G \right\|_F^2 + \lambda \left\| \sum_{i=1}^{q} L_i X L_i' \right\|_F^2 \right\}, \tag{4}$$

where the matrices $L_i, L_i' \in \mathbb{R}^{n\times n}$ are regularization matrices and $\lambda > 0$ is a regularization parameter; see, e.g., [1, 3, 4, 7, 18] for discussions on Tikhonov regularization. The minimization problem (4) has a unique solution $X = X_\lambda$ for any $\lambda > 0$ under suitable conditions on the regularization matrices L_i and L_i'. It is straightforward to analyze the unique solvability of (4) by transforming the problem to a minimization problem for the vector $x = \mathrm{vec}(X)$; cf. (2).

Our solution method is applicable when the matrices A_i, B_i, L_i, L_i' are square and of the same size. No other structure is required, because the matrices are only used in the evaluation of matrix-matrix products. Moreover, the size-restriction can be circumvented by zero-padding. We note that many popular regularization matrices are finite difference approximations of a derivative and are not square; see Section 4 for an illustration. Several approaches to construct square regularization matrices are discussed in [16].

It is the purpose of the present paper to describe a structure-preserving iterative method for the solution of (4) based on a generalization of the global Arnoldi method. The latter scheme, which was introduced in [11], is a block Arnoldi method that requires many fewer inner product evaluations than the standard block Arnoldi method; see also [2] for properties of the global Arnoldi method. The reduction in arithmetic operations is particularly significant when the block size is large. This is the case in the image deblurring application considered in this paper. A detailed comparison of the arithmetic work required by different block Arnoldi methods is provided in [11] and timings are reported in [1]. Our generalization of the global Arnoldi method described in [11] is analogous to the generalization of the standard Arnoldi process presented by Li and Ye [13]. The latter scheme has been applied to the solution of linear discrete ill-posed problems in [15].

We remark that image deblurring problems of the form (4) arise when the point-spread function is separable. This is, for instance, the case for Gaussian blur. Examples can be found in Section 4. Moreover, many blurring matrices can be approximated well by a sum of the form (1) with p small; see the discussion by Kamm and Nagy [12] on how a large matrix can be approximated by a Kronecker product, a topic also is treated in [14, 19].

This paper is organized as follows. Section 2 presents a generalized global Arnoldi method for a matrix pair $\{A, L\}$ and shows some properties of this method. An application to the solution of the Tikhonov minimization problem (4) is discussed in Section 3, a few computed examples are described in Section 4, and concluding remarks can be found in Section 5.

We conclude this section by introducing some notation. The Frobenius inner product is given by

$$\langle V, W \rangle_F = \mathrm{trace}(V^T W),$$

where V and W are matrices of appropriate sizes. Then $\|V\|_F = \langle V, V\rangle_F^{1/2}$. The matrices V and W are said to be F-orthogonal if $\langle V, W\rangle_F = 0$. Moreover, the sequence of matrices $V_1, V_2, V_3, \ldots$ is said to be F-orthonormal if

$$\langle V_j, V_k \rangle_F = \begin{cases} 0 & j \neq k, \\ 1 & j = k. \end{cases}$$

Above and throughout this paper, matrices are treated as "vectors" in the linear space $\mathbb{R}^{n\times n}$. We also will use the standard inner product

$$\langle f, g\rangle_2 = f^T g, \qquad f, g \in \mathbb{R}^r$$

and Euclidean vector norm

$$\|f\|_2 = \langle f, f\rangle_2^{1/2}, \qquad f \in \mathbb{R}^r.$$

2. Generalized global Arnoldi methods

2.1. A generalized global Arnoldi method for matrix pairs

Let $A, L \in \mathbb{R}^{n\times n}$ be large, possibly sparse or structured, matrices. Application of k steps of the generalized global Arnoldi method to the matrix pair $\{A, L\}$ with initial matrix $V \in \mathbb{R}^{n\times s}$, where $1 \leq s \ll n^2$, determines the matrix

$$\mathcal{V}(:, 1:(2k+1)s) = [V_1, V_2, \ldots, V_{2k+1}], \tag{5}$$

whose F-orthonormal blocks $V_j \in \mathbb{R}^{n\times s}$ form a basis for the $(2k+1)s$-dimensional generalized matrix Krylov subspace spanned by the first $2k+1$ of the blocks

$$V, AV, LV, A^2V, ALV, LAV, L^2V, \ldots .$$

The generalized global Arnoldi process also determines matrices H_A and H_L with k columns and about $2k+1$ rows; see below for details. The choice of k is commented on in Section 4.

In (5) and below, we use MATLAB-like notation: $\mathcal{V}(:, 1 : js)$ denotes a submatrix that is made up of all rows and columns 1 through js of the matrix $\mathcal{V}$. We refer to the scheme implemented by Algorithm 1 below as the generalized global Arnoldi process, because it generalizes the global Arnoldi process discussed in [1, 2, 11]. When $s = 1$, the algorithm reduces to the generalized Arnoldi process introduced by Li and Ye [13].

Algorithm 1. The generalized global Arnoldi process for the matrix pair $\{A, L\}$ and initial matrix V:

```
1.  V_1 := V/||V||_F; N := 1;
2.  for j = 1,...,k do
    2.1. if j > N then exit;
    2.2. Ṽ := A V_j ;
    2.3. for i = 1,...,N do
              H_A(i,j) := <Ṽ, V_i>_F;
              Ṽ := Ṽ - H_A(i,j) V_i ;
    2.4. end for
    2.5. H_A(N+1,j) := ||Ṽ||_F;
    2.6. if H_A(N+1,j) > 0 then
              N := N+1;
              V_N := Ṽ/H_A(N,j)
         else
              exit;
    2.7. end if
    2.8. Ṽ := L V_j;
    2.9. for i = 1,...,N do
              H_L(i,j) := <Ṽ, V_i>_F;
              Ṽ := Ṽ - H_L(i,j) V_i ;
    2.10. end for
    2.11. H_L(N+1,j) := ||Ṽ||_F;
    2.12. if H_L(N+1,j) > 0 then
              N := N+1;
              V_N := Ṽ/H_L(N,j)
          else
              exit;
    2.13. end if
3.  end for
```

Let α_k and β_k be the values of the number N at the end of lines 2.7 and 2.13 of Algorithm 1, respectively, at the last iteration (when $j = k$). We obtain from Algorithm 1 the relations

$$
\begin{aligned}
A\mathcal{V}(:,1:ks) &= \mathcal{V}(:,1:\alpha_k s)\,(H_{A,k}\otimes I_s),\\
L\mathcal{V}(:,1:ks) &= \mathcal{V}(:,1:\beta_k s)\,(H_{L,k}\otimes I_s),
\end{aligned}
$$

where $H_{A,k} = H_A(1:\alpha_k, 1:k)$ and $H_{L,k} = H_L(1:\beta_k, 1:k)$. The iterations with Algorithm 1 are terminated in case of breakdown, i.e., when the inequalities in lines 2.6 or 2.12 are violated.

2.2. A generalized global Arnoldi method for linear operator pairs

Consider the two linear operators

$$
\begin{aligned}
\mathcal{A}:\mathbb{R}^{n\times n} &\longrightarrow \mathbb{R}^{n\times n}\\
X &\longrightarrow \mathcal{A}(X)=\sum_{i=1}^{p} A_i\,X\,B_i
\end{aligned}
$$

and

$$
\begin{aligned}
\mathcal{L}:\mathbb{R}^{n\times n} &\longrightarrow \mathbb{R}^{n\times n}\\
X &\longrightarrow \mathcal{L}(X)=\sum_{i=1}^{q} L_i\,X\,L_i',
\end{aligned}
$$

which are applied in (4). The generalized global Arnoldi algorithm for the solution of (4) is deduced from Algorithm 1 by replacing lines 2.2 and 2.8 by

$$\tilde{V} := \mathcal{A}(V_j) \tag{6}$$

and

$$\tilde{V} := \mathcal{L}(V_j), \tag{7}$$

respectively, and by setting $s = n$. It is not difficult to show the relations

$$[\mathcal{A}(V_1),\mathcal{A}(V_2),\ldots,\mathcal{A}(V_k)] = \mathcal{V}(:,1:\alpha_k n)\,(H_{A,k}\otimes I_n), \tag{8}$$

$$[\mathcal{L}(V_1),\mathcal{L}(V_2),\ldots,\mathcal{L}(V_k)] = \mathcal{V}(:,1:\beta_k n)\,(H_{L,k}\otimes I_n) \tag{9}$$

for the algorithm so obtained. The following section applies this algorithm to the solution of (4).

3. Solution of the Tikhonov regularization problem

Introduce the subspaces

$$\mathbb{E}_k = \mathrm{span}\{V_1, V_2, \ldots, V_k\}, \quad k = 1, 2 \ldots\ ,$$

and define

$$\mathcal{V}_r = \mathcal{V}(:, 1 : rn) = [V_1, V_2, \ldots, V_r],$$

where $V_1, V_2, \ldots, V_r$ are F-orthonormal matrices generated by Algorithm 1 with the expressions (6) and (7) replacing those of lines 2.2 and 2.8, respectively, and with initial block V_1. After k steps of the algorithm, we determine an approximate solution $X_k \in \mathbb{E}_k$ of (4) of the form

$$X_k = \sum_{i=1}^{k} y_k^{(i)} V_i = \mathcal{V}_k\,(y_k \otimes I_n), \tag{10}$$

where $y_k^{(i)}$ is the ith component of the vector $y_k \in \mathbb{R}^k$. We need the following result to describe the computation of X_k.

Proposition 1. *The F-orthogonal projection $\widetilde{G}$ of the matrix G onto the space $\mathbb{E}_r$ is given by $\widetilde{G} = \mathcal{V}_r(z_G \otimes I_n)$ with $z_G = (\langle V_1, G\rangle_F, \ldots, \langle V_r, G\rangle_F)^T$. For all $z, g \in \mathbb{R}^r$, we have*

1. $\langle \mathcal{V}_r(z \otimes I_n), \mathcal{V}_r(g \otimes I_n)\rangle_F = \langle z, g\rangle_2$ *and* $\|\mathcal{V}_r(z \otimes I_n)\|_F = \|z\|_2$.
2. $\langle \mathcal{V}_r(z \otimes I_n), G\rangle_F = \langle z, z_G\rangle_2$.
3. $\|G - \widetilde{G}\|_F^2 = \|G\|_F^2 - \|z_G\|_2^2 \geq 0$.
4. $\|\mathcal{V}_r(z \otimes I_n) - G\|_F^2 = \|z - z_G\|_2^2 + \|G\|_F^2 - \|z_G\|_2^2$.

Proof It is clear that $\widetilde{G} \in \mathbb{E}_r$ and since the block vectors $\{V_1, \ldots, V_r\}$ are F-orthonormal, we have

$$\langle \widetilde{G} - G, V_i\rangle_F = 0, \quad i = 1, 2, \ldots, r.$$

Property 1 follows from the fact that

$$\langle \mathcal{V}_r(z \otimes I_n), \mathcal{V}_r(g \otimes I_n)\rangle_F = \sum_{i,j=1}^{r} \langle z_i V_i, g_j V_j\rangle_F = \sum_{i=1}^{r} z_i g_i = \langle z, g\rangle_2$$

for all $z = [z_i] \in \mathbb{R}^r$ and $g = [g_i] \in \mathbb{R}^r$.

Property 2 is a consequence of the relation

$$\langle \mathcal{V}_r(z \otimes I_n), G\rangle_F = \langle \sum_{i=1}^{r} z_i V_i, G\rangle_F = \sum_{i=1}^{r} z_i \langle V_i, G\rangle_F = \langle z, z_G\rangle_2$$

for all $z = [z_i] \in \mathbb{R}^r$.

Finally, properties 3 and 4 are immediately obtained from the F-orthogonal projection properties. □

Requiring the solution of the Tikhonov regularization problem (4) to live in $\mathbb{E}_k$ yields the minimization problem

$$\min_{X_k \in \mathbb{E}_k} \{ \, \|\mathcal{A}(X_k) - G\|_F^2 + \lambda \, \|\mathcal{L}(X_k)\|_F^2 \, \}. \tag{11}$$

Using the representation (10) of X_k, we obtain

$$\mathcal{A}(X_k) = \sum_{i=1}^{k} y_k^{(i)} \mathcal{A}(V_i) \quad \text{and} \quad \mathcal{L}(X_k) = \sum_{i=1}^{k} y_k^{(i)} \mathcal{L}(V_i),$$

which also can be written as

$$\mathcal{A}(X_k) = [\mathcal{A}(V_1), \mathcal{A}(V_2), \ldots, \mathcal{A}(V_k)] \, (y_k \otimes I_n) \tag{12}$$

and

$$\mathcal{L}(X_k) = [\mathcal{L}(V_1), \mathcal{L}(V_2), \ldots, \mathcal{L}(V_k)] \, (y_k \otimes I_n). \tag{13}$$

Substituting (8) and (9) into (12) and (13), respectively, gives

$$\mathcal{A}(X_k) = \mathcal{V}_{\alpha_k} \, (H_{A,k} \otimes I_n) \, (y_k \otimes I_n) = \mathcal{V}_{\alpha_k} \, (H_{A,k} \, y_k \otimes I_n) \tag{14}$$

and

$$\mathcal{L}(X_k) = \mathcal{V}_{\beta_k} \, (H_{L,k} \otimes I_n) \, (y_k \otimes I_n) = \mathcal{V}_{\beta_k} \, (H_{L,k} \, y_k \otimes I_n). \tag{15}$$

The relations (14) and (15) in conjunction with Proposition 1 show that

$$\begin{aligned} \|\mathcal{A}(X_k) - G\|_F^2 &= \left\|\mathcal{V}_{\alpha_k} \left(H_{A,k} \, y_k \otimes I_n\right) - G\right\|_F^2 \\ &= \left\|H_{A,k} \, y_k - z_G\right\|_2^2 + \|G\|_F^2 - \|z_G\|_2^2 \end{aligned}$$

and

$$\|\mathcal{L}(X_k)\|_F = \left\|\mathcal{V}_{\beta_k} \, (H_{L,k} \, y_k \otimes I_n)\right\|_F = \left\|H_{L,k} \, y_k\right\|_2.$$

Consequently, the minimization problem (11) can be expressed as the low-dimensional Tikhonov regularization problem

$$\min_{y_k \in \mathbb{R}^k} \{ \|H_{A,k}\, y_k - z_G\|_2^2 + \lambda \|H_{L,k}\, y_k\|_2^2 \}. \tag{16}$$

We illustrate the use of the initial blocks

$$V_1 = \frac{\mathcal{A}(G)}{\|\mathcal{A}(G)\|_F} \quad \text{and} \quad V_1 = \frac{G}{\|G\|_F} \tag{17}$$

in the numerical examples of Section 4.

It is important to use a suitable value of the regularization parameter λ in Tikhonov regularization (4) and (16). A variety of techniques are available for determining such a value. We will describe the use of generalized cross-validation (GCV) and the L-curve criterion. Since the size of the reduced Tikhonov regularization problem (16) is small, the computations required for determining a suitable value of λ are not very demanding.

We first discuss the application of GCV; see [4, 5]. Here λ is chosen to minimize the GCV function

$$\begin{aligned} GCV(\lambda) &= \frac{||H_{A,k} y_{k,\lambda} - z_G||_2^2}{[\text{trace}(I_{\alpha_k} - H_{A,k}\mathcal{H}_{\lambda,k}^{-1} H_{A,k}^T)]^2} \\ &= \frac{||(I_{\alpha_k} - H_{A,k}\mathcal{H}_{\lambda,k}^{-1} H_{A,k}^T) z_G||_2^2}{[\text{trace}(I_{\alpha_k} - H_{A,k}\mathcal{H}_{\lambda,k}^{-1} H_{A,k}^T)]^2}, \end{aligned} \tag{18}$$

where $y_{k,\lambda}$ is the solution of (16) and

$$\mathcal{H}_{\lambda,k} = H_{A,k}^T H_{A,k} + \lambda H_{L,k}^T H_{L,k}.$$

It is convenient to express (18) with the aid of the generalized singular value decomposition (GSVD) of the matrix pair $\{H_{A,k}, H_{L,k}\}$ given by

$$\begin{aligned} H_{A,k} &= U_k\, C_k\, M_k, \\ H_{L,k} &= V_k\, S_k\, M_k, \end{aligned}$$

where $U_k \in \mathbb{R}^{\alpha_k \times \alpha_k}$ and $V_k \in \mathbb{R}^{\beta_k \times \beta_k}$ are orthogonal matrices,

$$C_k = \text{diag}[c_1, c_2, \ldots, c_k] \in \mathbb{R}^{\alpha_k \times k}, \quad S_k = \text{diag}[s_1, s_2, \ldots, s_k] \in \mathbb{R}^{\beta_k \times k}$$

are diagonal matrices with nonnegative diagonal entries such that $c_i^2+s_i^2=1$ for all i, and $M_k \in \mathbb{R}^{k\times k}$ is a nonsingular matrix; see, e.g., [6] for details. The GCV function can be written as

$$GCV(\lambda) = \frac{\lambda^2 \sum_{i=1}^{k} \left(\frac{s_i^2 b_i}{c_i^2+\lambda s_i^2} \right)^2 + \sum_{i=k+1}^{\alpha_k} {b_i}^2}{\left(\sum_{i=1}^{k} \frac{c_i^2}{c_i^2+\lambda s_i^2} - \alpha_k \right)^2}$$

with $b=[b_1,\ldots,b_{\alpha_k}]^T := U_k^T z_G$.

Alternatively, we may avoid the computation of the GSVD by first determining the QR factorization

$$H_{L,k} = Q_{L,k}\, R_{L,k},$$

where $Q_{L,k} \in \mathbb{R}^{\beta_k\times k}$ has orthonormal columns and $R_{L,k} \in \mathbb{R}^{k\times k}$ is upper triangular. Typically, $R_{L,k}$ is nonsingular and not very ill-conditioned. We will assume this to be the case. Then setting $z_k = R_{L,k} y_k$, the Tikhonov minimization problem (16) reduces to the simpler minimization problem

$$\min_{z_k\in\mathbb{R}^k} \{ \|\tilde{H}_{A,k}\, z_k - z_G\|_2^2 + \lambda \|z_k\|_2^2 \}, \tag{19}$$

where

$$\tilde{H}_{A,k} = H_{A,k}\, R_{L,k}^{-1}. \tag{20}$$

The GCV function for the problem (19) is

$$\begin{aligned} GCV(\lambda) &= \frac{||\tilde{H}_{A,k} z_{k,\lambda} - z_G||_2^2}{[\mathrm{trace}(I - \tilde{H}_{A,k}\mathcal{K}_{\lambda,k}^{-1}\tilde{H}_{A,k}^T)]^2} \\ &= \frac{||(I_{\alpha_k} - \tilde{H}_{A,k}\mathcal{K}_{\lambda,k}^{-1}\tilde{H}_{A,k}^T) z_G||_2^2}{[\mathrm{trace}(I_{\alpha_k} - \tilde{H}_{A,k}\mathcal{K}_{\lambda,k}^{-1}\tilde{H}_{A,k}^T)]^2}, \end{aligned} \tag{21}$$

where $z_{k,\lambda}$ is the solution of the problem (19) given by

$$z_{k,\lambda} = \mathcal{K}_{\lambda,k}^{-1}\, \tilde{H}_{A,k}^T\, z_G, \quad \text{and} \quad \mathcal{K}_{\lambda,k} = \tilde{H}_{A,k}^T\, \tilde{H}_{A,k} + \lambda\, I_k.$$

Consider the singular value decomposition (SVD)

$$\tilde{H}_{A,k} = {U_k}' S_k' (V_k')^T, \tag{22}$$

where $U_k{}' \in \mathbb{R}^{\alpha_k \times \alpha_k}$ and $V_k' \in \mathbb{R}^{k \times k}$ are orthonormal matrices. The matrix

$$S_k' = \mathrm{diag}[s_1', s_2', \ldots, s_k'] \in \mathbb{R}^{\alpha_k \times k}$$

contains the singular values $s_1' \geq s_2' \geq \ldots \geq s_k' \geq 0$. From (21), we obtain the expression

$$GCV(\lambda) = \frac{\lambda^2 \sum_{i=1}^{k} \left(\frac{\tilde{b}_i}{s_i'^2 + \lambda} \right)^2 + \sum_{i=k+1}^{\alpha_k} \tilde{b}_i^2}{\left(\sum_{i=1}^{k} \frac{s_i'^2}{s_i'^2 + \lambda} - \alpha_k \right)^2}, \tag{23}$$

where $\tilde{b} = [\tilde{b}_1, \ldots, \tilde{b}_{\alpha_k}]^T := (U_k{}')^T z_G$. We used (23) in the computed examples reported in Section 4.

The L-curve criterion [8, 9] is another popular method for determining a suitable value of λ. Consider the graph

$$\lambda \longrightarrow \{||H_{A,k}\, y_{k,\lambda} - z_G||_2, ||H_{L,k}\, y_{k,\lambda}||_2\}.$$

This graph often has the shape of the letter "L" at least in a neighborhood of the desired λ-value and is referred to as the L-curve. The L-curve criterion prescribes that the value of λ that corresponds to the "vertex" of the "L" be chosen.

It is attractive to express the L-curve in terms of quantities of the minimization problem (19). Thus, the L-curve is the graph

$$\lambda \longrightarrow \{||\tilde{H}_{A,k}\, z_{k,\lambda} - z_G||_2, ||z_{k,\lambda}||_2\},$$

where $\tilde{H}_{A,k}$ is defined by (20). The singular value decomposition (22) of $\tilde{H}_{A,k}$ yields

$$\|\tilde{H}_{A,k}\, z_\lambda - z_G\|_2^2 = \sum_{i=1}^{k} \left(\frac{\lambda \tilde{b}_i}{s_i'^2 + \lambda} \right)^2 + \sum_{i=k+1}^{\alpha_k} \tilde{b}_i^2, \tag{24}$$

$$\|z_{k,\lambda}\|_2^2 = \sum_{i=1}^{k} \left(\frac{s_i' \tilde{b}_i}{s_i'^2 + \lambda} \right)^2. \tag{25}$$

Suppose that the "vertex" of the L-curve is associated with a value smaller than $\lambda_{\max}$ of the parameter λ. We then allocate N points,

$$0 = \lambda_1 < \lambda_2 < \ldots < \lambda_{N-1} < \lambda_N = \lambda_{\max},$$

in the interval $[0, \lambda_{\max}]$, and use the formulas (24)-(25) to determine, for each iteration k, the points

$$P_{i,k} = \{\|\tilde{H}_{A,k}\, z_{\lambda_i} - z_G\|, \|z_{\lambda_i}\|\}, \qquad i = 1, 2, \ldots, N, \tag{26}$$

on the L-curve. These points are employed to compute the approximate location of the vertex of the L-curve by using a slight modification of the algorithm by Rodriguez and Theis [17, p. 80]. Our implementation does not use the logarithm of the coordinates of the points (26). We found this modification to perform better for the problems considered in this paper. In the examples of Section 4, we let $\lambda_{\max} = 500$ and $N = 5000$, however, smaller values also can be used.

When the "vertex" of the L-curve is determined by visual inspection, often the logarithm of the coordinates of the points $P_{i,k}$, defined by (26), are plotted instead of the points $P_{i,k}$. We plot the $P_{i,k}$ in the next section, because the graphs are not used to locate the "vertex."

4. Numerical examples

This section presents a few numerical examples concerned with the solution of linear discrete ill-posed problems of the form (1) with a right-hand side matrix G that is contaminated by an error E. We illustrate the performance of our method in the context of image deblurring. All computations were carried out using MATLAB version 6.5 on an Intel Pentium workstation with about 16 significant decimal digits.

The exact (blur- and noise-free) gray scale image is denoted by $\widehat{X}$ in all examples. It is represented by an array of $n \times n$ pixels with values in the range $[0, 255]$ and is assumed not to be available. The matrix

$$\widehat{G} = \mathcal{A}(\widehat{X}) = \sum_{i=1}^{p} A_i\, \widehat{X}\, B_i,$$

represents an unavailable blurred, but noise-free, image associated with $\widehat{X}$. Here A_i and B_i are blurring matrices. Finally, the matrix (3) represents the available blur- and noise-contaminated image that we would like to restore. The noise-matrix E has normally distributed random entries with zero mean and with variance chosen so that

$$||E||_F/||\widehat{G}||_F = 10^{-\eta} \tag{27}$$

for some specified value of $\eta > 0$.

We show the exact, contaminated, and restored images. This provides a qualitative measure of the restored images X_k. Quantitative measures are provided by the relative error

$$e_k = \frac{||\widehat{X} - X_k||_F}{||\widehat{X}||_F}$$

and the peak signal-to-noise ratio (PSNR)

$$\mathrm{PSNR}(\widehat{X}, X_k) = 10\,\log_{10}\Big(\frac{n^2 d^2}{||\widehat{X} - X_k||_F^2}\Big).$$

The parameter d is 255, the largest pixel value.

Example 4.1. We let the exact (blur- and noise-free) image be the `enamel` image from MATLAB. It is represented by an array of 256×256 pixels and is shown on the left-hand side of Figure 1. In this example, we let $p = 1$ and $q = 1$ in (4). The blurring matrices of size 256×256 are the uniform Toeplitz matrix $A_1 = [a_{ij}]$ given by

$$a_{ij} = \begin{cases} \frac{1}{2r-1}, & |i-j| \leq r, \\ 0, & \text{otherwise}, \end{cases}$$

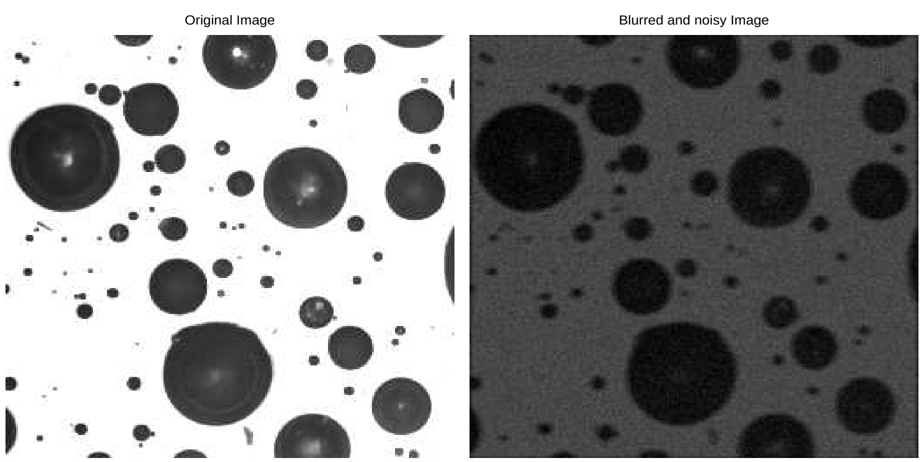

Figure 1: Exact image (left) and contaminated image (right).

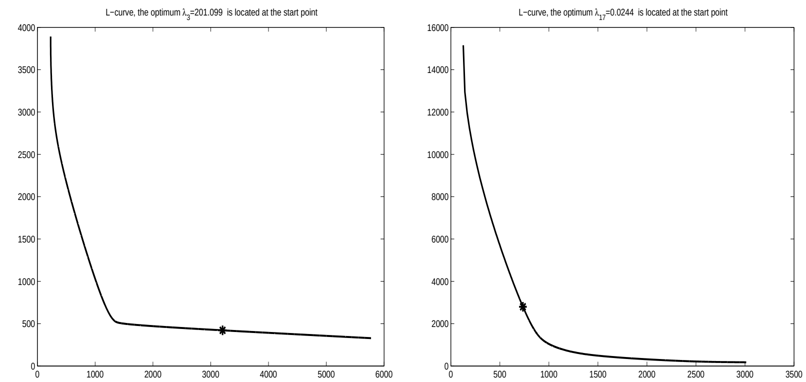

Figure 2: L-curves with the location of the chosen "vertices", which correspond to $\lambda_3 \simeq 201.99$ at iteration $k = 3$ (left) and $\lambda_{17} \simeq 0.0244$ at iteration $k = 17$ (right).

and the Gaussian Toeplitz matrix $B_1 = [b_{ij}]$ defined by

$$b_{ij} = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(i-j)^2}{2\sigma^2}\right), & |i-j| \leq r, \\ 0, & \text{otherwise.} \end{cases}$$

The variance is $\sigma = 7$, and $r = 2$ for both matrices. The matrix A_1 models out-of-focus blur and B_1 models atmospheric blur. The noise satisfies (27)

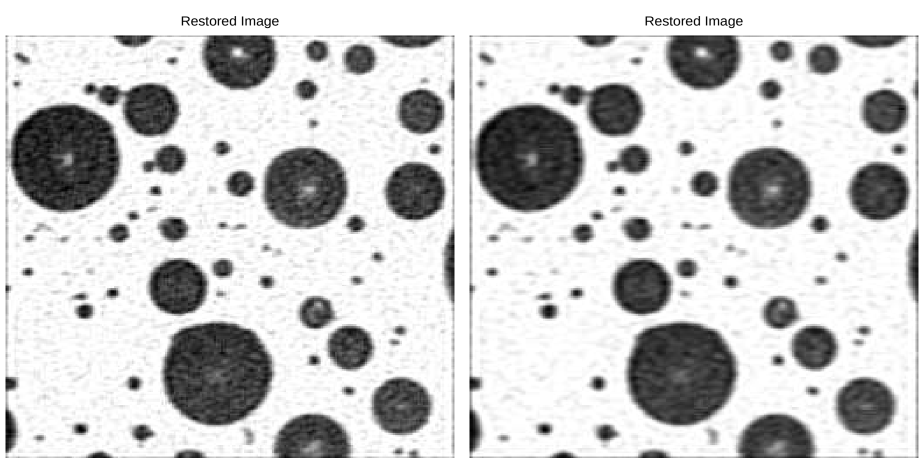

Figure 3: Restored images with $V_1 = \frac{G}{||G||_F}$ (left) and with $V_1 = \frac{\mathcal{A}(G)}{||\mathcal{A}(G)||_F}$ (right).

with $\eta = 0.95$. The PSNR-value for the blur- and noise-contaminated image G, i.e., PSNR$(\widehat{X}, G)$, is 4.17. This image is shown on the right-hand side of Figure 1.

We use the regularization matrices

$$L_1 = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 1 & -1 \\ 0 & \dots & \dots & 0 & 0 \end{bmatrix} \in \mathbb{R}^{256\times 256}, \qquad L_1' = L_1^T.$$

These matrices represent discrete first order derivatives.

k	$V_1 = \frac{G}{\|G\|_F}$		$V_1 = \frac{\mathcal{A}(G)}{\|\mathcal{A}(G)\|_F}$	
	Relative error e_k	PSNR	Relative error e_k	PSNR
1	9.734e-001	16.331	1.180e-001	19.956
5	1.232e-001	19.583	1.100e-001	20.567
15	9.859e-002	21.522	9.698e-002	21.664
20	9.739e-002	21.628	9.650e-002	21.708

Table 1: Relative errors and PSNR-values for a few restorations X_k.

We applied the L-curve criterion to determine a suitable value of the regularization parameter λ at each iteration with Algorithm 1. Figure 2 shows the L-curves at iterations $k = 3$ and $k = 17$ and the "vertex" determined by the algorithm in [17]. The vertices correspond to the parameter values $\lambda_3 \simeq 201.099$ and $\lambda_{17} \simeq 0.0244$ after 3 and 17 iterations, respectively. The fact that $\lambda_3 \gg \lambda_{17}$ indicates that three iteration steps are insufficient. Indeed, the choice of k may be based on the behavior of the regularization parameter values λ_k. For instance, we may increase k until the values λ_k do not vary much with k.

Table 1 displays values of the relative error e_k and PSNR-values for restorations X_k determined after k iterations for a few values of k. The initial blocks for Algorithm 1 are given by (17). The relative errors e_k are seen to be smaller and the PSNR-values larger for the initial block $V_1 = \mathcal{A}(G)/\|\mathcal{A}(G)\|_F$ than for the initial block $V_1 = G/\|G\|_F$. This is in agreement with Figure 3, which shows restored images obtained with these initial blocks. □

Example 4.2. The exact (blur- and noise-free) image `fruit` is represented by an array of 512×512 pixels. It is shown on the left-hand side of Figure 4. In this example, we let $p = 1$ and $q = 1$ in (4). The Gaussian Toeplitz blurring matrix $A_1 \in \mathbb{R}^{512\times512}$ has the entries

$$a_{ij} = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(i-j)^2}{2\sigma^2}\right), & |i-j| \le r, \\ 0, & \text{otherwise,} \end{cases}$$

Figure 4: Exact image (left) and contaminated image (right).

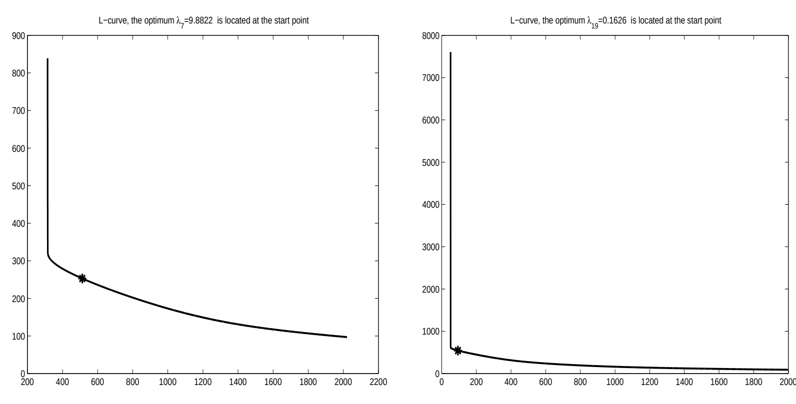

Figure 5: L-curves with the location of the "vertices", which correspond to $\lambda_7 \simeq 9.882$ at the iteration $k = 7$ (left) and $\lambda_{19} \simeq 0.162$ at the iteration $k = 19$ (right).

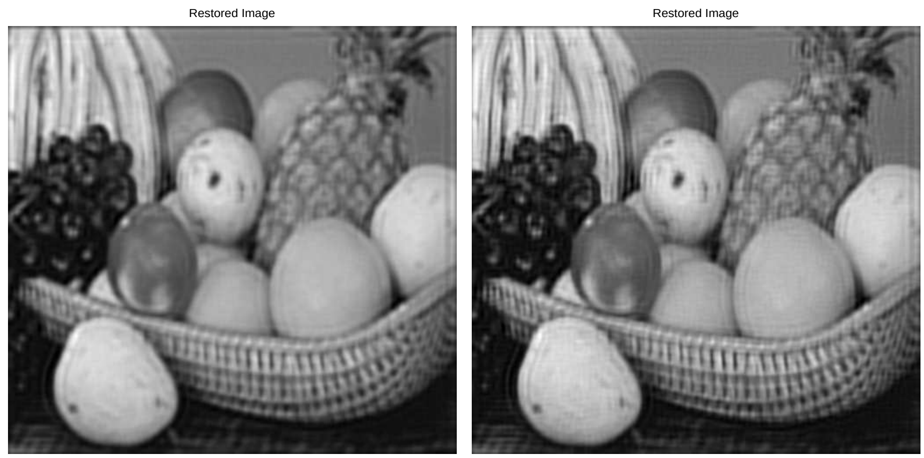

Figure 6: Restored images with $V_1 = \frac{G}{||G||_F}$ (left) and with $V_1 = \frac{\mathcal{A}(G)}{||\mathcal{A}(G)||_F}$ (right).

with $\sigma = 5$ and $r = 5$. We let $B_1 = A_1$. The noise satisfies (27) with $\eta = 2$. The PSNR-value of the contaminated image G, i.e., PSNR$(\widehat{X}, G)$, is 11.99. The image G is shown on the right-hand side of Figure 4. The regularization matrices L_1 and L_1' are of size 512×512 and of the same kind as in Example 4.1.

k	$V_1 = \frac{G}{\|\|G\|\|_F}$		$V_1 = \frac{\mathcal{A}(G)}{\|\|\mathcal{A}(G)\|\|_F}$	
	Relative error e_k	PSNR	Relative error e_k	PSNR
1	1.309e-001	23.394	1.512e-001	22.148
5	1.220e-001	24.006	1.086e-001	25.020
15	1.082e-001	25.052	9.845e-002	25.874
20	9.539e-002	26.149	8.746e-002	26.903

Table 2: Relative errors and PSNR-values for a few restorations X_k.

We applied the L-curve criterion to determine a suitable value of the regularization parameter λ for the projected problem at each iteration. Figure 5 shows the the L-curves at iterations $k = 7$ and $k = 19$. They gave the parameter values $\lambda_7 \simeq 9.882$ and $\lambda_{19} \simeq 0.162$, respectively.

Table 2 shows relative errors e_k and PSNR-values for a few iterates X_k for the initial blocks $V_1 = G/\|G\|_F$ and $V_1 = \mathcal{A}(G)/\|\mathcal{A}(G)\|_F$. The table shows the latter to yield the best restorations. This is also illustrated by Figure 6. □

Example 4.3. In this experiment, we consider the case $p = 2$ and $q = 1$ in (4) with

$$\mathcal{A}(X) = A_1 X B_1 + A_2 X B_2.$$

Here A_1 and B_1 are the same matrices as in Example 4.2 with $\sigma = 5$ and $r = 5$. The matrices A_2 and B_2 are of the same kind as the matrices A_1 and B_1 in Example 4.1 with $\sigma = 7$ and $r = 3$. The regularization matrices L_1 and L_1' are the same as in Example 4.2.

The exact (blur- and noise-free) image, `fruit`, is the same as in Example 4.2. It is shown on the left-hand side of Figure 4. The noise satisfies (27) with $\eta = 2$. The PSNR-value for the blur- and noise contaminated image G is 22.373. This image is shown on the left-hand side of Figure 7. In this example, we applied the GCV criterion to determine a suitable value of the regularization parameter λ at each iteration with Algorithm 1, starting with the initial block $V_1 = \mathcal{A}(G)/\|\mathcal{A}(G)\|_F$. Table 3 shows the PSNR-value to improve from 22.373 to 28.598. The restored image with the latter PSNR-value is displayed on the right-hand side of Figure 7. □

k	Relative error e_k	PSNR	λ_k with GCV
1	1.387e-001	22.892	10.00
5	9.163e-002	26.499	1.458
15	8.035e-002	27.639	0.126
20	7.196e-002	28.598	0.002

Table 3: Relative errors, PSNR-values and the optimal values of the regularized parameter λ_k (with GCV).

5. Conclusion

We derived a generalized global Arnoldi method, which is a block method analogous to the single-vector scheme described in [13], and discussed its application to Tikhonov regularization of linear discrete ill-posed problems with a structured matrix. Image restoration examples are presented.

Figure 7: Contaminated image (left) and restored image (right).

Acknowledgment

We would like to thank the referees for comments that improved the presentation.

References

[1] A. Bouhamidi and K. Jbilou, *Sylvester Tikhonov-regularization methods in image restoration*, J. Comput. Appl. Math., 206 (2007), pp. 86–98.

[2] L. Elbouyahyaout, A. Messaoudi, and H. Sadok, *Algebraic properties of the block GMRES and block Arnoldi methods*, Electron. Trans. Numer. Anal., 33 (2009), pp. 207–220.

[3] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

[4] G. H. Golub and U. von Matt, *Tikhonov regularization for large scale problems*, in: G. H. Golub, S. H. Lui, F. Luk, and R. Plemmons (Eds.), Workshop on Scientific Computing, Springer, New York, 1997, pp. 3–26.

[5] G. H. Golub, M. Heath, and G. Wahba, *Generalized cross-validation as a method for choosing a good ridge parameter*, Technometrics 21 (1979), pp. 215–223.

[6] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.

[7] C. W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, 1984.

[8] P. C. Hansen *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Rev., 34 (1992), pp. 561–580.

[9] P. C. Hansen and D. P. O'Leary, *The use of the L-curve in the regularization of discrete Ill-posed problems*, SIAM J. Sci. Comput., 14 (1993), pp. 1487–1503.

[10] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[11] K. Jbilou, A. Messaoudi, and H. Sadok, *Global FOM and GMRES algorithms for matrix equations*, Appl. Numer. Math., 31 (1999), pp. 49–63.

[12] J. Kamm and J. G. Nagy, *Kronecker product and SVD approximations in image restoration*, Linear Algebra Appl., 284 (1998), pp. 177–192.

[13] R.-C. Li and Q. Ye, *A Krylov subspace method for quadratic matrix polynomials with application to constrained least squares problems*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 405–428.

[14] N. P. Pitsianis, *The Kronecker Product in Approximation and Fast Transform Generation*, Ph.D. thesis, Department of Computer Science, Cornell University, 1997.

[15] L. Reichel, F. Sgallari, and Q. Ye, *Tikhonov regularization based on generalized Krylov subspace methods*, Appl. Numer. Math., to appear.

[16] L. Reichel and Q. Ye, *Simple square smoothing regularization operators*, Electron. Trans. Numer. Anal., 33 (2009), pp. 63–83.

[17] G. Rodriguez and D. Theis, *An algorithm for estimating the optimal regularization parameter by the L-curve*, Rend. Mat. Appl., 25 (2005), pp. 69–84.

[18] A. N. Tikhonov, *Regularization of incorrectly posed problems*, Soviet Math. Dokl., 4 (1963), pp. 1624–1627.

[19] C. Van Loan and N. Pitsianis, *Approximation with Kronecker products*, in: M. S. Moonen, G. H. Golub, and B. L. R. De Moor (Eds.), Linear Algebra for Large Scale and Real-Time Applications, Kluwer, Dordrecht, 1993, pp. 293–314.