DOI 10.7603/s40836-016-0009-z

GSTF Journal of Mathematics, Statistics and Operations Research (JMSOR) Vol.3 No.2, July 2016

GKB-GCV Method for Solving Generic Tikhonov Regularization Problems

Dai Togashi
Keio University
Graduate School of Science and Technology

3-14-1 Hiyoshi, Kohoku, Yokohama, JAPAN, 223-8522

Takashi Nodera
Keio University
Department of Mathematics

3-14-1 Hiyoshi, Kohoku, Yokohama, JAPAN, 223-8522
e-mail: nodera@math.keio.ac.jp

***Abstract*—The W-GCV method is one of the iterative methods used to solve a large scale standard form of the Tikhonov regularization, but it is also necessary for solving the general form of the Tikhonov regularization. This paper proposes a new solver, called GKB-GCV, which is an extension of the W-GCV by using the GSVD. Numerical results are presented to show the usefulness of the GKB-GCV method in large scale ill-posed problems.**

Keywords-GKB, W-GCV, Tikhonov regularization, large scale ill-posed problem, Image resolution.

I. INTRODUCTION

The stable approximate solution for a large scale ill-posed problem of the form:

$$\boldsymbol{x}_{\rm LS} = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\rm argmin}\, ||\boldsymbol{b} - A\boldsymbol{x}||_2^2 \qquad (1)$$

is computed, where matrix $A \in \mathbb{R}^{m\times n}$, $m \geq n$, is ill-conditioned. The right-hand vector $\boldsymbol{b} \in \mathbb{R}^m$ contains the following error:

$$\boldsymbol{b} = A\boldsymbol{x}_{\rm exact} + \boldsymbol{\epsilon}, \qquad (2)$$

where $x_{\rm exact} \in \mathbb{R}^n$ is the exact solution, and $\boldsymbol{\epsilon} \in \mathbb{R}^m$ is the unknown noise. A matrix of this form sometimes comes from image resolutions. Because matrix Ǝ is ill-conditioned, $\boldsymbol{x}_{\rm LS}$ is dependent on noise. The Tikhonov regularization [11] constructs stable approximations of $x_{\rm exact}$ by solving the least squares problem of the form:

$$\boldsymbol{x}_\lambda = \underset{\boldsymbol{x}\in\mathbb{R}^n}{\rm argmin}\{||\boldsymbol{b} - A\boldsymbol{x}||_2^2 + \lambda||L\boldsymbol{x}||_2^2\} \qquad (3)$$

where $L \in \mathbb{R}^{p\times n}$ is the regularization matrix, and $\lambda > 0$ is the regularization parameter. The standard form of the Tikhonov regularization is when $L = I_n$, where I_n is the $n \times n$ identity matrix. The general form of the Tikhonov regularization is when $L \neq I_n$. When the common space between the null spaces of Ǝ and L is the zero space, the regularization problem (3) has a unique solution. To obtain a good approximate solution for (3), an appropriate regularization parameter is required. There are many methods for determining the regularization parameter without knowledge of the noise's norm $||\boldsymbol{\epsilon}||_2$, [1, 5, 6].

Solving equation (3) and all the approach selecting parameters of the above is computable when the GSVD for the pair of matrix (Ǝ, L) has been computed. One problem is that the GSVD is not cost effective when the Ǝ or the L is a large scale matrix. For a large scale problem, iterative methods are used, e.g. the LSQR, the CGLS, or some kind of Krylov subspace method. Computing good approximate solutions by using iterative methods, require the parameter λ a priori and a suitable stopping criteria. The hybrid method solves this issue through combining the projection method with an inner regularization method. For $L = I_n$, there are two hybrid methods, called GKB-FP [2] and W-GCV [4]. These methods do not require identifying the norm $||\boldsymbol{\epsilon}||_2$, and contain a projection over the Krylov subspace generated by the Golub-Kahan Bidiagonalization (GKB) method. The difference between these two methods is in the approach in terms of determining the regularization parameter. The GKB-FP uses the FP scheme, whereas the W-GCV uses the weighed GCV.

Lampe et al. [8] and Reichel et al. [10] have proposed approaches for $L \neq I_n$ by minimizing the regularization problem over the generalized Krylov subspace. These approaches determine the regularization parameter by using the knowledge of the norm $||\boldsymbol{\epsilon}||_2$. Bazán et al. [3] proposed an approach without identifying the norm $||\boldsymbol{\epsilon}||_2$, which is created by the extension of the GKB-FP method.

This paper focuses on the W-GCV method which is a solver for a large scale standard form of the Tikhonov regularization which does not require identifying the norm of the noise. This paper proposes applying an extension of the W-GCV to the general form of the Tikhonov regularization. The approach of the W-GCV is based on the idea of the GKB-FP and the idea of the AT-GCV [9]. The stopping criteria of the W-GCV and the AT-GCV are also compared.

This paper is organized as follows: After the introduction, Section II summarizes the framework of the classical W-GCV method. In Section III, the extensions of the W-GCV to the general form of the Tikhonov regularization, are described briefly. Following this, a new scheme of GKB-GCV is proposed. In section IV, the usefulness of the GKB-GCV for test problems, is illustrated. The conclusions and possible future studies are explored in Section V.

DOI : 10.5176/2251-3388-3.2.71

Published online: 29 December 2016

II. The W-GCV Method

The W-GCV is one of the algorithms for the standard form of the Tikhonov regularization, which is based on the GKB and weighted GCV. For the standard form, i.e. $L = I_n$, the SVD of matrix A reduces the solution for equation (3) as follows:

$$\boldsymbol{x}_\lambda = \sum_{i=1}^{n} \frac{\sigma_i \boldsymbol{u}_i^T \boldsymbol{b}}{\sigma_i^2 + \lambda^2} \boldsymbol{v}_i \quad \text{with} \quad A = U\Sigma V^T$$

where each $U = [\boldsymbol{u}_1, ..., \boldsymbol{u}_m]$, $V = [\boldsymbol{v}_1, ..., \boldsymbol{v}_n]$ has left and right singular vectors of A, and Σ is the singular value of matrix A diagonal with $\sigma_1 \geq ... \geq \sigma_n \geq 0$, and 0 on the nondiagonal. When we apply $k < n$ GKB steps to matrix A with the initial vector $\boldsymbol{b}$, it results in two matrices $Y_{k+1} = [\boldsymbol{y}_1, ..., \boldsymbol{y}_{k+1}] \in \mathbb{R}^{m\times(k+1)}$ and $W_k = [\boldsymbol{w}_1, ..., \boldsymbol{w}_k] \in \mathbb{R}^{n\times k}$ with orthonormal columns, and a lower bidiagonal matrix as follows.

$$B_k = \begin{pmatrix} \alpha_1 & & & \\ \beta_2 & \alpha_2 & & \\ & \beta_3 & \ddots & \\ & & \ddots & \alpha_k \\ & & & \beta_{k+1} \end{pmatrix} \in \mathbb{R}^{(k+1)\times k},$$

$$\begin{aligned} \beta_1 Y_{k+1} \boldsymbol{e}_1 &= \boldsymbol{b} = \beta_1 \boldsymbol{y}_1, \\ AW_k &= Y_{k+1} B_k, \\ A^T Y_{k+1} &= W_k B_k^T + \alpha_{k+1} \boldsymbol{w}_{k+1} \boldsymbol{e}_{k+1}^T, \end{aligned}$$

where $\boldsymbol{e}_i$ denotes the *i*-th unit vector in $\mathbb{R}^{k+1}$. Furthermore, columns of W_k are the orthonormal basis for the generalized Krylov subspace $\mathcal{K}_k(A^TA,\ A^T\boldsymbol{b})$. The standard form of regularization, i.e. $L = I_n$, over the generated Krylov subspace is as follows:

$$\boldsymbol{x}_\lambda^{(k)} = \underset{\boldsymbol{x}\in\mathcal{K}_k(A^TA,\ A^T\boldsymbol{b})}{\operatorname{argmin}} \{||A\boldsymbol{x} - \boldsymbol{b}||_2^2 + \lambda ||\boldsymbol{x}||_2^2\}. \quad (4)$$

Since the columns of W_k are the orthonormal basis for the generated Krylov subspace, equation (4) is reduced as follows:

$$\begin{aligned} \boldsymbol{x}_\lambda^{(k)} &= W_k \boldsymbol{y}_\lambda^{(k)}, \\ \boldsymbol{y}_\lambda^{(k)} &= \underset{\boldsymbol{y}\in\mathbb{R}^k}{\operatorname{argmin}} \{||B_k \boldsymbol{y} - \beta_1 \boldsymbol{e}_1||_2^2 + \lambda ||\boldsymbol{y}||_2^2\}. \end{aligned} \quad (5)$$

This reduction technique is a good choice for large scale problems, because this approach reduces the size of the least squares problem: $(m + p) \times n$ to $(2k + 1) \times k$.

The GCV and weighted GCV methods determine the regularization parameter. The GCV determines the regularization parameter by searching for the minimum point of function as follows:

$$G_{A,b,I_n}(\lambda) = \frac{||(I_m - AA_{\lambda,I_n}^+)b||_2^2}{(\text{trace}(I_m - AA_{\lambda,I_n}^+))^2}, \quad (6)$$

where $A_{\lambda,L}^+ = (A^TA + \lambda^2 L^T L)^{-1} A^T$. Using the SVD for matrix A, equation (6) is written as follows:

$$G_{A,\boldsymbol{b},I_n}(\lambda) = \frac{\sum_{i=1}^{n} \left(\frac{\lambda^2 \boldsymbol{u}_i^T \boldsymbol{b}}{\sigma_i^2 + \lambda^2}\right)^2 + \sum_{i=n+1}^{m} (\boldsymbol{u}_i^T \boldsymbol{b})^2}{\left(m - \sum_{i=1}^{n} \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}\right)^2}. \quad (7)$$

The approach of the GCV to the least squares problem is as follows:

$$G_{B_k,\beta_1\boldsymbol{e}_1,I_k}(\lambda) = \frac{||(I_{k+1} - B_k(B_k)_{\lambda,I_k}^+)\beta_1 \boldsymbol{e}_1||_2^2}{(\text{trace}(I_{k+1} - B_k(B_k)_{\lambda,I_k}^+))^2}. \quad (8)$$

However, the optimal parameter determined by equation (8) is unsuitable for equation (3). Therefore, the weighted GCV for a reduced system of equation (5) is used instead:

$$G_{B_k,\beta_1\boldsymbol{e}_1,I_k}^{(\omega)}(\lambda) = \frac{||(I_{k+1} - B_k(B_k)_{\lambda,I_k}^+)\beta_1 \boldsymbol{e}_1||_2^2}{(\text{trace}(I_{k+1} - \omega B_k(B_k)_{\lambda,I_k}^+))^2}. \quad (9)$$

When $\omega = 1$, the weighted GCV is the same as the standard GCV method. Furthermore, the approximate solution becomes smooth at $\omega > 1$, and less smooth at $\omega < 1$. Similarly, the SVD for matrix B_k reduces equation (9) as follows:

$$G_{B_k,\beta_1\boldsymbol{e}_1,I_k}^{(\omega)}(\lambda) = \frac{\beta_1^2 \left(\sum_{i=1}^{k} \left(\frac{\lambda \boldsymbol{u}_{i(k)}^T \boldsymbol{e}_1}{\sigma_{i(k)}^2 + \lambda}\right)^2 + (\boldsymbol{u}_{k+1(k)}^T \boldsymbol{e}_1)^2\right)}{\left(k + 1 - \omega \sum_{i=1}^{k} \frac{\sigma_{i(k)}^2}{\sigma_{i(k)}^2 + \lambda}\right)^2}. \quad (10)$$

where $B_k = U_k \Sigma_k V_k^T$. The SVD for B_k can be computed easily, because the size of B_k is $(k+1) \times k$, and smaller than the size of matrix A. The stopping criterion is as follows:

$$\frac{|\hat{G}(k+1) - \hat{G}(k)|}{|\hat{G}(1)|} < tol \quad (11)$$

where $\hat{G}(k)$ is an approximation for $G_{A,\boldsymbol{b},I_n}(\lambda)$ without a weighted parameter, and *tol* is the stopping tolerance.

$$\begin{aligned} \hat{G}(k) &= \frac{||(I_m - AW_k(B_k)_{\lambda,I}^+ Y_{k+1}^T)\boldsymbol{b}||_2^2}{(\text{trace}(I_m - AW_k(B_k)_{\lambda,I}^+ Y_{k+1}^T))^2}, \\ &= \frac{\beta_1^2 \left(\sum_{i=1}^{k} \left(\frac{\lambda_k^2 \boldsymbol{u}_{i(k)}^T \boldsymbol{e}_1}{\sigma_{i(k)}^2 + \lambda_k^2}\right)^2 + (\boldsymbol{u}_{k+1(k)}^T \boldsymbol{e}_1)^2\right)}{\left(m - \sum_{i=1}^{k} \frac{\sigma_{i(k)}^2}{\sigma_{i(k)}^2 + \lambda_k^2}\right)^2}. \end{aligned}$$

The W-GCV method is summarized in Algorithm 1.

Algorithm 1 W-GCV

Require: $A, \boldsymbol{b}, \omega > 0, tol$

Ensure: Regularized solution $\boldsymbol{x}_{\lambda^*}^{(k)}$

1. Apply GKB step to A with starting vector $\boldsymbol{b}$ at $k = 0$ and set $k = 1$.
2. Perform one more GKB step.
3. Compute SVD(B_k). $B_k = U_k \Sigma_k V_k^T$.
4. Compute λ_k by weighted GCV.
5. **If** stopping criteria is satisfied **do**
   $\lambda^* = \lambda_k$.
   else do
   $k \leftarrow k + 1$
   Update the weighted parameter ω
   Go to step 2.
   end if
6. Solve subproblem $\boldsymbol{y}_{\lambda^*}^{(k)}$.
7. Compute the regularized solution $\boldsymbol{x}_{\lambda^*}^{(k)}$.

III. GKB-GCV METHOD

The purpose of this section is to identify a working extension of the W-GCV that can be used with a general form of the Tikhonov regularization. The extension that is proposed will be referred to as the GKB-GCV. In its general form, i.e. $L \neq I_n$, equation (3) is reduced by the GKB as follows:

$$\boldsymbol{x}_\lambda^{(k)} = W_k \boldsymbol{y}_\lambda^{(k)},$$
$$\boldsymbol{y}_\lambda^{(k)} = \underset{\boldsymbol{y}\in\mathbb{R}^k}{\operatorname{argmin}}\{||B_k\boldsymbol{y} - \beta_1\boldsymbol{e}_1||_2^2 + \lambda||LW_k\boldsymbol{y}||_2^2\}. \quad (12)$$

In equation (12), the size of the least squares problem is $(k+1+p)\times k$.

The same reduction to PROJ-L when solving the general form of the Tikhonov regularization of Bazán [3] is used as follows:

$$\boldsymbol{y}_\lambda^{(k)} = \underset{\boldsymbol{y}\in\mathbb{R}^k}{\operatorname{argmin}}\{||B_k\boldsymbol{y} - \beta_1\boldsymbol{e}_1||_2^2 + \lambda||R_k\boldsymbol{y}||_2^2\}, \quad (13)$$

where $LW_k = Q_k R_k$, using the QR factorization. For increasing k, the QR factorization can be updated computing $k+1$ elements by using the summation and a product of the vectors. This approach can be used without limitation of dimension for L, i.e. for any number of p, unlike the AT-GCV method which is one of the hybrid methods using the same GCV. The next step was to consider change points in the GCV. One problem is that when $L \neq I_n$, the SVD for matrix $\exists$ can not reduce the number of the residual norm and trace into the GCV function to form at equation (7) and (10). This problem was addressed by using the GSVD for the pair of matrix ($\exists$, L), $A = USZ^{-1}$and $L = VCZ^{-1}$, where $U = [\boldsymbol{u}_1, ..., \boldsymbol{u}_m]$, $V = [\boldsymbol{v}_1, ..., \boldsymbol{v}_p]$ are orthogonal, Z is nonsingular matrix , and each S, C have $s_1 \geq ... \geq s_n \geq 0$ and $0 \leq c_1 \leq ... \leq c_n$ on its diagonals and 0 on its nondiagonals. GSVD for the pair of matrix ($\exists$, L) reduces the GCV function as follows:

$$G_{A,\boldsymbol{b},L}(\lambda) = \frac{\sum_{i=1}^n \left(\frac{c_i^2\lambda\boldsymbol{u}_i^T\boldsymbol{b}}{s_i^2+c_i^2\lambda}\right)^2 + \sum_{i=n+1}^m (\boldsymbol{u}_i^T\boldsymbol{b})^2}{\left(m - \sum_{i=1}^n \frac{s_i^2}{s_i^2+c_i^2\lambda}\right)^2}. \quad (14)$$

Using a similar computation of $G^{(\omega)}_{B_k,\beta_1\boldsymbol{e}_1,R_k}(\lambda)$ the equation (14) was reduced as follows by using GSVD(B_k, R_k):

$$G^{(\omega)}_{B_k,\beta_1\boldsymbol{e}_1,R_k}(\lambda) = \frac{\beta_1^2\left(\sum_{i=1}^k \left(\frac{c_{i(k)}^2\lambda\boldsymbol{u}_{i(k)}^T\boldsymbol{e}_1}{s_{i,(k)}^2+c_{i(k)}^2\lambda}\right)^2 + (\boldsymbol{u}_{k+1(k)}^T\boldsymbol{e}_1)^2\right)}{\left(k+1-\omega\sum_{i=1}^k \frac{s_{i(k)}^2}{s_{i(k)}^2+c_{i(k)}^2\lambda}\right)^2}.$$

where $B_k = U_k S_k Z_k^{-1}$, $R_k = V_k C_k Z_k^{-1}$. However, the determination of weight parameter Z is difficult. The AT-GCV method applies a similar function to $\hat{G}(k)$ for the GCV function at step k [9]:

$$G_k(\lambda) = \frac{||(I_m - AW_k(B_k)^+_{\lambda,R_k}Y_{k+1}^T)\boldsymbol{b}||_2^2}{(\operatorname{trace}(I_m - AW_k(B_k)^+_{\lambda,R_k}Y_{k+1}^T))^2},$$
$$= \frac{\beta_1^2\left(\sum_{i=1}^k\left(\frac{\lambda_k^2\boldsymbol{u}_{i(k)}^T\boldsymbol{e}_1}{\sigma_{i(k)}^2+\lambda_k^2}\right)^2 + (\boldsymbol{u}_{k+1(k)}^T\boldsymbol{e}_1)^2\right)}{\left(m - \sum_{i=1}^k\frac{\sigma_{i(k)}^2}{\sigma_{i(k)}^2+\lambda_k^2}\right)^2}.$$

This approach does not need to determine weight parameter Z. These two functions have different purposes. The GCV function in W-GCV determines the appropriate regularization parameter for the reduced equation (13), and the GCV function at AT-GCV approximates the appropriate regularization parameter for the original equation (3). Furthermore, the AT-GCV uses the residual norm entered when computing the GCV function for the stopping rule which is different from the W-GCV:

$$\frac{|(||\boldsymbol{r}_{\lambda_k}^{(k)}||_2 - ||\boldsymbol{r}_{\lambda_{k-1}}^{(k-1)}||_2)|}{||\boldsymbol{r}_{\lambda_k}^{(k)}||_2} < \sqrt{tol}, \quad (15)$$

where $\boldsymbol{r}_{\lambda_k}^{(k)} = B_k\boldsymbol{y}_{\lambda_k}^{(k)} - \beta_1\boldsymbol{e}_1$. Numerical experiments were used to illustrate the differences between these stopping rules. The stopping rule for equation (11) to the $\hat{G}$) and the rule of equation (15) to GKB-GCV($r^{(k)}$) were noted. The stopping rule for GKB-GCV($r^{(k)}$) was too severe compared to the stopping rule for GKB-GCV($\hat{G}$). Hence, to create tolerance with regards to the stopping rule, $\sqrt{tol}$ on GKB-GCV($r^{(k)}$) was used. In addition, another stopping rule was used:

$$\frac{|(||\boldsymbol{r}_{\lambda_k}^{(k)}||_2 - ||\boldsymbol{r}_{\lambda_{k-1}}^{(k-1)}||_2)|}{||\boldsymbol{r}_{\lambda_1}^{(1)}||_2} < tol. \quad (16)$$

The GKB-GCV was compactly summarized in Algorithm 2.

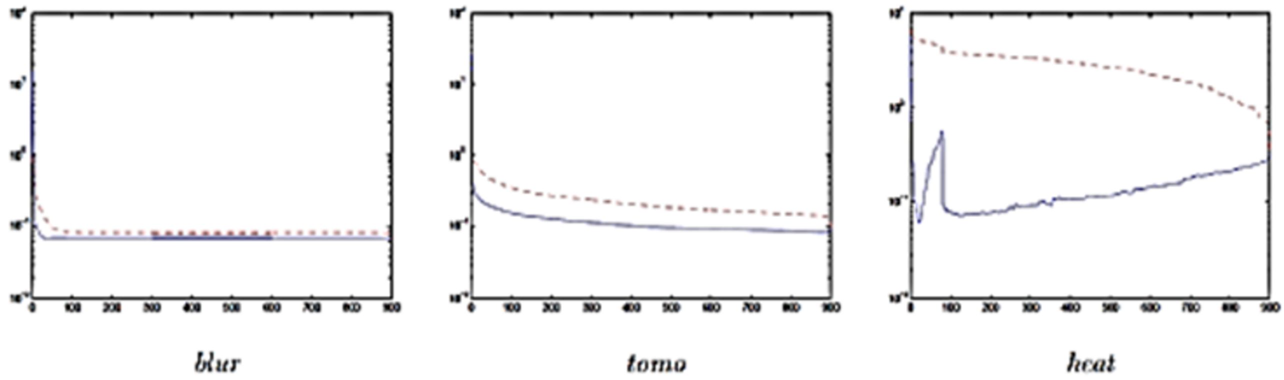

Figure 1: Change of relative error norm for the increase in k for the three test problem, *blur*, *tomo* and *heat*.

Algorithm 2 GKB-GCV

Require: $A, \boldsymbol{b}, L, tol$
Ensure: Regularized solution $\boldsymbol{x}_{\lambda^*}^{(k)}$

1. Apply GKB step to A with starting vector $\boldsymbol{b}$ at $k=0$ and set $k=1$.
2. Perform one more GKB step and update QR factorization of LW_k. $LW_k = Q_k R_k$.
3. Compute GSVD$(B_k,\ R_k)$. $B_k = U_k S_k Z^{-1},\ R_k = V_k C_k Z^{-1}$.
4. Compute minimized point λ_k of $G_k(\lambda)$.
5. **If** stopping criteria is satisfied **do**
   $\lambda^* = \lambda_k$.
   else do
   $k \leftarrow k+1$
   Go to step 2.
   end if
6. Solve subproblem $\boldsymbol{y}_{\lambda^*}^{(k)}$.
7. Compute the regularized solution $\boldsymbol{x}_{\lambda^*}^{(k)}$.

IV. Convergence Analysis of the GKB-GCV Method

In this section, we provide convergence properties of the GKB-GCV method. We define the appropriate parameter λ_* and λ_k as follows.

$$\lambda_* = \mathrm{argmin} G(\lambda),$$
$$\lambda_k = \mathrm{argmin} G_k(\lambda)$$

where $\lambda_n = \lambda_*$. Using triangle inequality, following inequality is satisfied.

$$||\boldsymbol{x}_{\text{exact}} - \boldsymbol{x}_{\lambda_k}^{(k)}||_2 \leq ||\boldsymbol{x}_{\text{exact}} - \boldsymbol{x}_{\lambda_n}^{(n)}||_2 + \sum_{i=k}^{n-1} ||\boldsymbol{x}_{\lambda_{i+1}}^{(i+1)} - \boldsymbol{x}_{\lambda_i}^{(i)}||_2 \quad (17)$$

The first term of equation (17) is corresponding to an error which occurs as a consequence of stabilization. The second term of (17) will converge monotonically for increasing *k*, and we verify its convergence property by following experiments.

A. Behavior of relative error norm at each GKB-FP iteration

We use built-in data in MATLAB, *blur, tomo* and *heat*, for test problem, and we use *n*=30 for *blur* and *tomo*, and *n*=900 for *heat*. In this time, we don't use stopping criteria and continue iteration until iteration number arrving at matrix size. In the figure.1 , solid lines represent the left-hand side of (17) and dashed lines represent the right-hand side of (17). For *blur* and *tomo*, we could bound well, but the right-hand side of (17) is too larger than the left-hand side of (17) for *heat*. The reason why the right-hand side of (17) is too large is that for the regularizaiton, relartive error norm rise or fall down after arriving at minimum relative error norm.

V. Numerical Experiments

The PROJ-L method which is also one of the hybrid methods and an extension of GKB-FP, was used for the purpose of comparision with the proposed method in this paper. The 2D image deblurring problem which is the procedure for recovering original images from blurred images using noise from the form equation (2) was considered. Matrix A was the blurring operator, e.g. the Point Spread Function(PSF) matrix, and $\boldsymbol{b}_{\text{exact}} = A\boldsymbol{x}_{\text{exact}}$ are blurred images without any noise. All computations of numerical experiments were carried out in MATLAB R2013b, and generated noise vectors $\boldsymbol{\epsilon}$ by the MATLAB code **randn** and $\text{NL} = ||\boldsymbol{\epsilon}||_2/||\boldsymbol{b}_{\text{exact}}||_2$. $N \times N$ images with $N^2 \times N^2$ Gaussian PSF matrices as a blurring operator were used. The Gaussian PSF was defined by $A = (2\pi\sigma^2)^{-1} T \otimes T$, where ς was the parameter used to control the width of the Gaussian PSF, and T was an $N \times N$ symmetric banded toeplitz matrix with generators of the form:

```
z= [exp(-([0:band-1].^2)/(2*σ^2));zeros(1,N-band)]
```

from Kilmore et al. [7]. In our tests, we used ς = 2 and **band** = 16. A regularization matrix lowering the gap between adjacent points was chosen:

$$L = \begin{bmatrix} I_N \otimes L_1 \\ L_1 \otimes I_N \end{bmatrix}, \quad L_1 = \begin{bmatrix} 1 & -1 & & \\ & 1 & -1 & \\ & & \ddots & \ddots \\ & & 1 & -1 \end{bmatrix} \in \mathbb{R}^{(N-1)\times N}.$$

unlike the other methods, the relative error did not decrease with the noise level. The relative error of the GKB-GCV($\hat{G}$) was about 1.5 times as more than the PROJ-L when the noise level was small.

	PROJ-L($p_0 = 10$)	GKB-GCV($\hat{G}$)	GKB-GCV($r^{(k)}$)	GKB-GCV($r^{(1)}$)
$\mathrm{NL} = 10^{-2}$				
$\bar{\lambda}$	0.0819	6.82×10^{-3}	5.57×10^{-3}	3.25×10^{-3}
$\bar{E}$	0.0821	0.0899	0.0848	0.0806
$\bar{t}$(sec)	0.416	0.150	0.206	0.399
$k_m(k_M)$	27(32)	11(11)	15(15)	27(27)
$\mathrm{NL} = 10^{-3}$				
$\bar{\lambda}$	7.11×10^{-3}	3.22×10^{-3}	4.90×10^{-4}	5.43×10^{-4}
$\bar{E}$	0.0651	0.0891	0.0681	0.0693
$\bar{t}$ (sec)	1.45	0.150	0.728	0.592
$k_m(k_M)$	94(102)	11(11)	47(48)	41(41)
$\mathrm{NL} = 10^{-4}$				
$\bar{\lambda}$	1.16×10^{-3}	3.15×10^{-3}	1.41×10^{-4}	2.24×10^{-4}
$\bar{E}$	0.0624	0.0891	0.0645	0.0675
$\bar{t}$ (sec)	2.03	0.150	1.41	0.734
$k_m(k_M)$	124(138)	11(11)	82(88)	49(53)

Table 1: Results for the test problem *rice*64 with $tol = 10^{-4}$

A. *Test problem 1: rice64*

The interpolation data of MATLAB, *rice* image were used in test problem 1. Firstly, a 64×64 sub-image of *rice* and *rice*64 were used to compare noise levels. $A \in \mathbb{R}^{4096\times 4096}$ and $L \in \mathbb{R}^{8064\times 4096}$, and the condition number was $\mathrm{cond}(A) \approx 2.14 \times 10^{16}$. These experiments used ten noise vectors for each noise level: $\mathrm{NL} = 10^{-2}, 10^{-3}$ and 10^{-4}. To simplify the notation, $\bar{\lambda}, \bar{t}$ and $\bar{E}$ denoted the average value of the regularization parameter, time and relative error, and $k_m(k_M)$ denoted the minimum (maximum) number of steps required.

The computation of the FP method on PROJ-L started with $p_0 = 10$ and $\Pi = 1$. The stopping criteria was set to $tol = 10^{-4}$.

All proposed methods converged faster than the existing method PROJ-L from Table 1. Previous experiments suggested to us that PROJ-L converged comparatively faster in all of the solvers for the general form of the Tikhonov regularization, and were dependent on the noise level. The GKB-GCV($r^{(k)}$) and the GKB-GCV($r^{(1)}$) also were dependent on noise level. The GKB-GCV($r^{(k)}$) had the same dependence as the PROJ-L. The dependence of the GKB-GCV($r^{(1)}$) was smaller than that of the GKB-GCV($r^{(k)}$). The results for the GKB-GCV($r^{(1)}$) were not much different for $\mathrm{NL} = 10^{-2}$ and 10^{-3}. Regarding numerical precision, all results of the proposed methods were worse than PROJ-L, except for the results of the GKB-GCV($r^{(1)}$) for $\mathrm{NL} = 10^{-2}$. Because the GKB-GCV($\hat{G}$) is independent of noise levels GKB-GCV($\hat{G}$) is independent of noise levels unlike the other methods, the relative error did not decrease with the noise level. The relative error of the GKB-GCV($\hat{G}$) was about 1.5 times as more than the PROJ-L when the noise level was small.

B. *Test problem 2: rice*

The next step was to create a 256 × 256 original image of *rice*. $A \in \mathbb{R}^{65536\times 65536}$, $L \in \mathbb{R}^{130560\times 65536}$ and $\mathrm{cond}(A) \approx 3.40 \times 10^{16}$. The same experiment was performed using $\mathrm{NL} = 10^{-2}$ and 10^{-3} this time. The same notation was used for the case of *rice*64. The computation of the FP method on PROJ-L started with $p_0 = 15$ and $\Pi = 1$, and $tol = 10^{-4}$, for the stopping criteria.

Similar results were obtained for *rice*64, with the exception of GKB-GCV($r^{(1)}$) for $\mathrm{NL} = 10^{-2}$. Results from the GKB-GCV($r^{(1)}$) for $\mathrm{NL} = 10^{-2}$ was slow with less numerical accuracy than the GKB-GCV($r^{(k)}$) and the PROJ-L. One of the reason for this, is that hybrid methods do not have properties of monotone convergence. The stopping rules ofhybrid methods must not be too severe or too easy. An easy approach for solving this problem was employed, which used two stopping rules. Specifically, both equations (15) and (16) were applied to the stopping rule. Please see Figure 1 for the original image, deblurred image, and resolution images.

VI. CONCLUSION

The GKB-GCV is a new solver for the general form of the Tikhonov regularization problem; it is based on the W-GCV. The GKB-GCV was compared to the PROJ-L and the GKB-GCV, using different stopping rules by using two image deblurring problems. The results of the numerical experiments showed that each of the proposed methods had good advantages. The GKB-GCV($\hat{G}$) was the fastest, although its numerical precision was the worst in all scenarios. This was because the GKB-GCV($\hat{G}$) does not depend on noise level.

	PROJ-L($p_0 = 15$)	GKB-GCV($\hat{G}$)	GKB-GCV($r^{(k)}$)	GKB-GCV($r^{(1)}$)
NL $= 10^{-2}$				
$\bar{\lambda}$	0.0874	1.22×10^{-3}	9.54×10^{-4}	5.37×10^{-4}
$\bar{E}$	0.0831	0.0928	0.0830	0.0860
$\bar{t}$ (sec)	3.38	0.906	1.90	4.89
$k_m(k_M)$	26(26)	7(7)	14(14)	36(38)
NL $= 10^{-4}$				
$\bar{\lambda}$	6.98×10^{-3}	6.82×10^{-4}	1.09×10^{-4}	1.09×10^{-4}
$\bar{E}$	0.0659	0.0925	0.0698	0.0697
$\bar{t}$ (sec)	21.0	0.925	6.18	6.25
$k_m(k_M)$	137(137)	7(7)	44(44)	45(45)

Table 2: Results for the test problem *rice* with $tol = 10^{-4}$

Secondly, the GKB-GCV($r^{(k)}$) was very fast, but had less accuracy compared to the PROJ-L for all noise levels, while it had the the same dependency on noise levels to PROJ-L. Lastly, the GKB-GCV($r^{(1)}$) had a smaller dependence on noise level than GKB-GCV($r^{(k)}$) and PROJ-L.

REFERENCES

[1] Bazán, F. S. V., ‘‘Fixed-point iterations in determining the Tikhonov regularization parameter,” Inverse Problems, Vol. 24, 035001, 2008.

[2] Bazán, F. S. V. and Borges, L. S., ‘‘GKB-FP: an algorithm for large- scale discrete ill-posed problems,” , BIT, Vol. 50, pp. 481–507, 2010.

[3] Bazán, F. S. V., Cunha, M. C. C. and Borges, L. S., ‘‘Extension of GKB-FP algorithm to large-scale general-form Tikhonov Numerregularization,” . Linear Algebra Appl., Vol. 21, pp. 316–339, 2014.

[4] Chung, J., Nagy, J. G. and O'Leary, D. P., ‘‘A weighted-GCV method for Lanczos-hybrid regularization,” ETNA, Vol. 28, pp. 149–167, 2008.

[5] Golub, G. H., Heath, M. and Wahba, G., ‘‘Generalized cross-validation as a method for choosing a good ridge parameter,” Technometrics, Vol. 21, pp. 215–222, 1979.

[6] Hansen, P. C. and O'Leary, D. P., ‘‘The use of the L-curve in the regularization of discrete ill-posed problems,” SIAM J. on Sci. Comput., Vol. 14, pp. 1487–1503, 1993.

[7] Kilmer, M. E., Hansen, P. C., Español, M. I., ‘‘A projection-based approach to generalform Tikhonov regularization,” SIAM J. on Sci. Comput., Vol. 29, pp. 315–330, 2007.

[8] Lampe, J., Reichel, L. and Voss, H., ‘‘Large-scale Tikhonov regularization via reduction by orthogonal projection,” Linear Algebra and its Applications, Vol. 436, pp. 2845–2865, 2012.

[9] NovatGCi, P., and Russo, M. R., ‘‘A GCV based Arnoldi-Tikhonov regularization method,” BIT Numer. Math., Vol. 54, pp. 501–521, 2014.

[10] Reichel, L., Sgallari, F. and Ye, Q., ‘‘Tikhonov regularization based on generalized Krylov subspace methods,” Appl. Numer. Math., Vol. 62, pp. 1215–1228, 2012.

[11] Tikhonov, A.N., ‘‘Solution of incorrectly formulated problems and the regularization method,” Soviet Mathematics Doklady, Vol. 4, pp. 1035–1038, 1963.

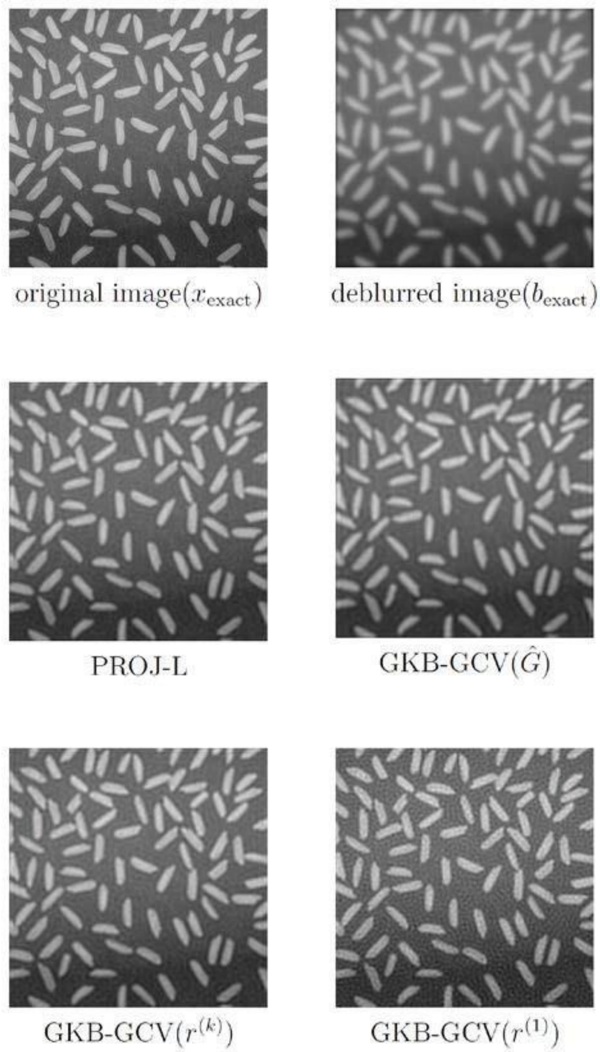

Figure 2: Resolution image of *rice* with NL $= 10^{-2}$