Regularised GMRES-type Methods for X-Ray Computed Tomography

Coban, S.B. and Lionheart, W.R.B.

2014

MIMS EPrint: **2015.31**

ISSN 1749-9097

Regularised GMRES-type Methods for X-Ray Computed Tomography

Sophia B. Coban and William R.B. Lionheart

***Abstract*—Slowly converging iterative methods such as Landweber or ART, have long been preferred for reconstructing a tomographic image from a set of CT data. In the recent years, a fast-converging method named CGLS has received attention for reconstructing tomographic data. However, there is a large class of methods that give more reliable solutions, when compared to CGLS. In this paper, we are going to consider the merits of the GMRES-type methods when applied to the CT problem, introduce various strategies, and compare the results with CGLS.**

In computed tomography, we deal with an over-determined system of linear equations of the form

$$Ax = b, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$ is a large and sparse geometry matrix, $b \in \mathbb{R}^{m\times 1}$ is the logarithm of the ratio of initial and final intensities, and $x \in \mathbb{R}^{n\times 1}$ is the linear attenuation coefficient in voxels. For computational efficiency and mathematical flexibility, iterative methods are preferred for solving problems of this type. Additionally, since the CT problem is very large, it cannot be solved with a direct method because that would require more matrices of the same size to be stored. This emphasises the need for iterative methods. In the CT community, when *iterative reconstruction methods* are discussed, one is accustomed to think of slowly converging algebraic iterative methods such as Landweber or ART. These methods have been around for a long time and are widely used because the user has the advantage of stopping the algorithm before the data is over-fit, even though the problem is mathematically not yet minimized. However, it is understood in the recent years that by reconstructing CT data with a fast-converging method, we also avoid over-fitting the data and are able to obtain a better solution. So in theory, to obtain an exact solution with these complex, fast-converging methods, we would have memory requirements that grow with the number of iterations but in practice, we require only a fraction of this number so our memory requirements stay low. These are just a few reasons why there has been a growing interest in fast-converging methods, with the most popular one being the **C**onjugate **G**radient method for **L**east **S**quares (CGLS). This method is mathematically equivalent to applying the original Conjugate Gradient to the normal equation, i.e.

$$A^T Ax = A^T b. \tag{2}$$

The performance of the CG method depends on the fact that the geometry matrix is symmetric, which automatically works well with (2) since the product of any matrix with its transpose will always be symmetric. However, it is not always ideal to calculate the exact transpose in CT problems, and it is much more efficient to implement an inexact transpose (also known as an unmatched back projection): A certain $\widehat{A^T}$ that is close, but not equal to A^T. Of course this means that the product of $\widehat{A^T}A$ is no longer a symmetric matrix, which causes some issues on the convergence and computational inefficiency with the CGLS method. This is an important point (to which we return to later on in the paper), and our main motivation for wanting to adapt alternative iterative methods in the same class as a well-established problem.

CGLS is a member of a large class of methods named the Krylov subspace (KS-)methods [8]. Those belonging to this class converge very quickly, which gives us the possibility of applying them directly to the Tikhonov system,

$$(A^T A + \lambda^2 L^T L)x = A^T b. \tag{3}$$

For the readers' convenience, we now give the definition of a Krylov subspace: An order k Krylov subspace, $\mathcal{K}_k(A, b)$, is the linear subspace spanned by the image of b under the linear transformation matrix A^p, $p = 0, 1, \ldots, k-1$ (where $A^0 = I_n$),

$$\mathcal{K}_k(A, b) = \text{span}\{b, Ab, A^2b, \ldots, A^{k-1}b\}. \tag{4}$$

KS-methods are derived from (4) and are popularly used for their convergence properties, robustness and efficiency. These methods are particularly preferred for when A is large and sparse since the product of Ab is a vector, and $A^2b = A(Ab)$ is another matrix-vector operation[1]. This avoids filling in the zero elements in the matrix and preserves the sparsity of A. Also, as $k \to \infty$, $A^k b \to A^{-1}b$, thus avoiding the inversion of a large and sparse matrix. Krylov subspace methods are also row (or column) action methods. This is important because in CT, the geometry matrix A is often too big to store, and the matrix-vector operations are required to be performed with one row (or column) of A at a time. So KS-methods are easily (and efficiently) adaptable for the CT problem.

Another popular method from the KS-methods class is GMRES. In the next section, we highlight the advantages and disadvantages of this method but we first give a quick introduction to GMRES and state the algorithm. We should

School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK.
www.maths.manchester.ac.uk/∼scoban, www.maths.manchester.ac.uk/∼bl.

[1]Note here that this definition is valid for square matrices. For non-square matrices, we deal with forward and back projection, rather than the powers of matrices.

note that a more detailed derivation of this method is given in [9], [5] as well as an extensive literature review. Here, we will mention only some of these references and briefly explain the GMRES variations.

GMRES-TYPE METHODS

The **G**eneralised **M**inimum **Res**idual (GMRES), is a KS-method that approximates a solution to (1) by evaluating

$$x_k = x_0 + V_k y. \tag{5}$$

Here, $V_k \in \mathbb{R}^{k\times k}$ is the orthonormal columns of basis for $\mathcal{K}_k(A, r_0)$, and $y \in \mathbb{R}^{k\times 1}$ is the solution to (what we refer to as) an *inner* problem, where $\|r_k\|_2$ is minimized over the Krylov subspace, $\mathcal{K}_k(A, r_0)$. The matrix V_k is obtained during what is called an *Arnoldi process*, which also returns a rectangular upper Hessenberg matrix, $H_k \in \mathbb{R}^{(k+1)\times k}$. This matrix mimics the characteristics of the coefficient matrix A and thus, it is used to obtain a solution to the inner problem, y, that minimizes the residual over $\mathcal{K}_k(A, r_0)$.

GMRES was first introduced by Saad and Schultz in 1986 [9], for solving nonsymmetric square matrices. Its convergence properties were studied by van der Horst, in 1993 [11], and its behaviour for singular and nearly singular matrices by Brown and Walker, in 1994 [1]. In the following years, there has been a great interest in the theory and applications of GMRES-type[2] methods: Complementing Brown and Walker's work, Calvetti proposed a GMRES-type method for singular matrices, called Range Restricted GMRES (RRGMRES) [2]. The idea, with which we experiment, was to shift $\mathcal{K}_k(A, r_0)$ by the coefficient matrix A prior to the Arnoldi process. The GMRES algorithm is given below:

ALGORITHM 1: *GMRES*
1. *Start*: Choose x_0. Let $r_0 = b - Ax_0$, $\beta = \|r_0\|_2$ and $V_1 = r_0/\beta$.
2. *Arnoldi Process*:
for $j = 1, 2, \ldots$ until convergence **do**
  $h_{(i,j)} = V_i^T A V_j$, $i = 1, 2, \ldots, j$,
  $\omega = AV_j - \sum_{i=1}^{j} h_{(i,j)} V_i$,
  $h_{(j+1,j} = \|\omega\|_2$,
  $V_{j+1} = \omega / h_{(j+1,j)}$.
end for
3. *Solve:* $\|\beta e_1 - H_{k+1,k} y\|_2$ for y, where $e_1 = [1, 0, \ldots, 0]^T$.
4. *Form the approximated solution:* $x_k = x_0 + V_k y$.

Step 3 can be done with the help of QR factorization coupled with Givens rotation. For the RRGMRES algorithm, we only have to replace $V_1 = r_0/\|r_0\|_2$ by $V_1 = Ar_0/\|Ar_0\|_2$ in Step 1.

GMRES is popularly used as a deblurring technique and often compared with other iterative methods [3]. It is also used as part of new hybrid methods [4] or coupled with preconditioners [13]. In recent years, GMRES is adapted to solve various other applied problems, e.g. inverse blackbody radiation problem [12], or non-rotational CT [10]. However most of these works are limited to square matrices and are concluded using simulated data.

CG is arguably the most popular KS-method because it is easy to implement, computationally inexpensive and numerically stable when applied to square, symmetric and positive-definite systems. As we mentioned earlier, CGLS is the equivalent of CG applied to $A^T Ax = A^T b$, where $A^T A$ is clearly symmetric but when an unmatched back projection, $\widehat{A}^T$, is used (where $\widehat{A}^T \neq A^T$), $\widehat{A}^T A$ is **not** symmetric. So CGLS has difficulties giving reliable results, whereas GMRES encounters no problems. This is because GMRES is designed to work with nonsymmetric systems as opposed to CG. Another advantage of GMRES is that, in case of a well-conditioned coefficient matrix in a square system (that is, $m = n$ in (1)), GMRES works as a direct solver and returns the exact solution in n steps. When compared to CGLS [3], examples show that the residual is much smaller and GMRES requires less computational work.

One disadvantage of GMRES is its large memory requirements. To avoid this, Saad and Schultz suggested to restart the method after a certain number of iterations (parameter m, chosen by user), clear the memory, and use the m^{th} iterate x_m as the new initial vector x_0, before the next cycle of iterations. In our experiments, we found that for normal matrices, restarted GMRES (or GMRES(m)) does not converge at all. Our results are omitted here but can be found in [5]. However, as we have already said, the applications to CT with typical accuracies of data, we do not need that many iterations before we are over-fitting. So the disadvantage of GMRES is not an issue for us when solving the CT problem.

In our strategies, we combine GMRES and RRGMRES with other useful tools (e.g. Tikhonov regularisation). Various test matrices from MATLAB's `gallery`, Hansen's `Regularization Tools` and `AIRTools` [6], [7] are used, but we include only two test cases with 1% and 10% Gaussian noise in this paper. The computations are carried out using MATLAB 2013a with a personal laptop of specifications 1.7 GHz Intel Core i7.

NUMERICAL EXPERIMENTS

In this section, we test the accuracy and speed of GMRES, RRGMRES and regularised GMRES methods. We have 8 strategies to compare and give results for two simulated cases: *Parallel 2D* and *FanBeam 2D*. The strategies are listed in Table I.

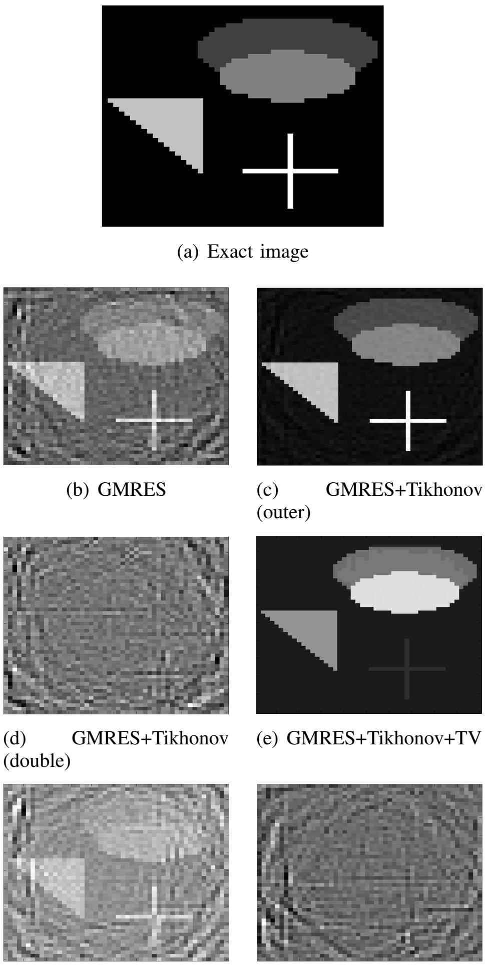
Test Case 1: Parallel 2D

This test case is generated using the `tomo(N,f)` function from Hansen's `Regularization Tools` [6]. N is chosen to be 50, i.e. matrix A has dimensions $N^2 \times N^2$. The images below are obtained for when b contains 1% noise, and the dataset is complete. For this and the next case, we assume we have no prior information about the object, so for the strategies involving Tikhonov, we take $L = I$ (and $\lambda = 10^{-4}$).

[2]Throughout the paper, we use this term to mean GMRES and its variations.

Strategy	Description
1) GMRES	$full$–GMRES algorithm, as stated in ALGORITHM 1.
2) GMRES+Tikhonov (outer)	GMRES algorithm applied to solve the Tikhonov system (3).
3) GMRES+Tikhonov (double)	GMRES algorithm applied to solve the Tikhonov system (3), and the *inner* problem in Step 3 is replaced by its Tikhonov alternative.
4) GMRES+Tikhonov +TV	GMRES algorithm applied to solve the Tikhonov system (3). The system is then plugged in TV to be solved (where the GMRES solution is used as the starting point).
5) RRGMRES	GMRES algorithm where $V_1 = r_0/\|r_0\|$ is replaced by $V_1 = Ar_0/\|Ar_0\|$ in Step 1.
6) RRGMRES+Tikhonov (outer)	RRGMRES algorithm applied to solve the Tikhonov system (3).
7) RRGMRES+Tikhonov (double)	RRGMRES algorithm applied to solve both the 'outer' Tikhonov system (3) and the Tikhonov alternative of the *inner* problem in Step 3.
8) CGLS	The popular CGLS algorithm, run until the same tolerance value is satisfied.

TABLE I
THE STRATEGIES USED IN THE TEST CASES, FOLLOWED BY THEIR DESCRIPTIONS.



(a) Exact image

(b) GMRES

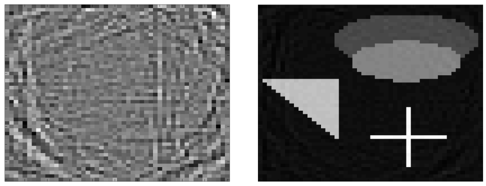
(c) GMRES+Tikhonov (outer)

(d) GMRES+Tikhonov (double)

(e) GMRES+Tikhonov+TV

(f) RRGMRES

(g) RRGMRES+Tikhonov (outer)



(h) RRGMRES+Tikhonov (double)

(i) CGLS (10^4 iterations)

Fig. 1. Results for *Parallel 2D* test problem with 1% noise.

We see that the features of the exact image are distinguishable in the GMRES runs. GMRES gives much better results when we start with a problem where Tikhonov regularisation is already added. The results seem to improve further when a TV solution is computed following GMRES+Tikhonov.

RRGMRES gives similar results to GMRES, but when coupled with Tikhonov, it fails to give any reasonable solutions. This is because RRGMRES is designed for singular or nearly singular systems, so it is not stable when the problem becomes 'less' ill-posed. We will not include RRGMRES in the future experiments in our work.

Finally, we note that all the GMRES runs took less than N^2 iterations whereas CGLS reached the maximum iteration number, which was set by us as 10^4.

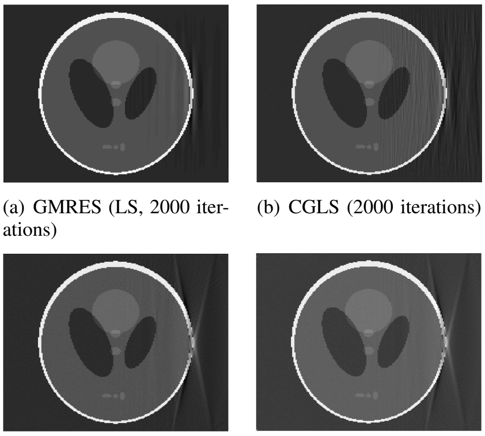
Test Case 2: FanBeam 2D

The second test case is generated by using the function `fanbeamtomo(N,angles,projections)` from Hansen's `AIRTools` Toolbox [7]. This time, N is chosen to be 200 (i.e. the image size is 200×200), the number of angles to be 180 (going from 0 to 179), and the number of projections to be 360. This means that the size of the geometry matrix, A, is $(180 \times 360) \times (200 \times 200) = 64800 \times 40000$. Additionally, we have added 10% Gaussian noise to the data vector to account for the experimental noise.

In the previous case we obtained some promising results with GMRES+Tikhonov (outer). However, for the strategies involving Tikhonov regularisation, the prior information was taken as $L = I$, and the regularisation parameter as $\lambda = 10^{-4}$. This means that starting GMRES with (3) is very close to starting GMRES with (2). So this time we run GMRES started with (2), and thus make a fairer comparison to CGLS (which is mathematically equivalent to CG started with (2)). We also compare these KS-methods to the popular iterative methods, Landweber and ART (the details of Landweber and ART algorithms can be found in [7]).

Fig. 2. Phantom image used for Test Case 2.

We run each method for 2000 iterations except for ART, which is not designed for large number of iterations. Comparing the solution norms with Landweber (at 2000^{th} iteration), we believe that running 40 iterations of ART is a fair comparison to running 2000 iterations of GMRES. The reconstructed images are presented in Fig. 3.



(a) GMRES (LS, 2000 iterations)

(b) CGLS (2000 iterations)

(c) Landweber (2000 iterations)

(d) ART (40 iterations)

Fig. 3. Results for *FanBeam 2D* test problem with 10% noise.

Since both GMRES and CGLS are members of the KS-methods class, it is not surprising to see both reconstructions with the same features: Both Fig. 3(a) and 3(b) have converged to the phantom image well, except for the second half. This is simply because we have not iterated enough. Interestingly, after iterating for a long time, Landweber or ART are still not as close to the phantom image as GMRES or CGLS are. In fact, the image reconstructed with Landweber at 2000 iterations can be obtained with 39 iterations of GMRES, and 51 iterations of CGLS. This highlights the benefits of fast-converging methods.

When compared to CGLS, although it is difficult to see from the images, the noise is somewhat less in Fig. 3(a), and the lines on the second half of the image are less pronounced.

Conclusion and Future Work

This paper briefly introduced the theory and algorithm of GMRES-type methods, as well as investigating the effects of combining GMRES and its variations with regularisation tools. A number of strategies were tested with simulated tomography data. The results achieved are promising and motivates a great number of possibilities, with which our implementation and reconstructions can be improved. They also provoke different ideas, which are all summarised and listed as future tasks below.

1) In our test cases, we have simulated a 2D parallel and fan beam experiments where 1% and 10% Gaussian noise was added to the tomographic data. However, to understand the benefits of regularised-GMRES, it is necessary that our strategies are tried with real datasets.
2) It is also necessary to apply these strategies to cases where some prior information is known and used in the Tikhonov system. It is important to see how this would affect the reconstructed images.
3) Our test cases showed GMRES + Tikhonov (outer) or GMRES (LS) can be used as alternatives to CGLS. An important next step could be testing these algorithms when an unmatched back projection is used.
4) In addition to that, one must apply these strategies to limited data problems, were the data is obtained with fewer angles.
5) The algorithms we discussed can also be further optimised and parallelised for the reconstruction of larger datasets or 3D and 4D (space + time) tomography.
6) We must also test these ideas against the popular reconstruction methods such as CGLS and FDK to highlight the advantages and disadvantages of our strategies. CGLS and FDK are available to users at our facilities in the University of Manchester.
7) More detailed investigation into GMRES with normal equations (rather than applying the algorithm to $A^T Ax = A^T b$) is needed. We need a clearer picture of how that can affect the convergence when there is noise in data.
8) Finally, one more task we can do is to make use of appropriate preconditioners in the GMRES strategies to improve the convergence properties (especially for the real data case).

Acknowledgement

This project is funded by BP through the ICAM framework.

References

[1] P.N. Brown and H.F. Walker. GMRES on (nearly) singular systems. *SIAM J. Matrix Anal. Appl.*, 18:37–51, 1994.

[2] D. Calvetti, B. Lewis, and L. Reichel. GMRES-type methods for inconsistent systems. *Linear Algebra Appl.*, 316:157–169, 2000.

[3] D. Calvetti, B. Lewis, and L. Reichel. On the choice of subspace for iterative methods for linear discrete ill-posed problems. *Int. J. Appl. Math. Comput. Sci.*, 1:1069–1092, 2001.

[4] D. Calvetti, B. Lewis, and L. Reichel. A hybrid GMRES and TV-norm based method for image restoration, 2002.

[5] S.B. Coban. Regularised GMRES-type methods for x-ray computed tomography. Technical report, School of Mathematics, University of Manchester, United Kingdom, September 2013. Available on request.

[6] P.C. Hansen. Regularisation tools version 4.0 for MATLAB 7.3. *Numer. Algorithms*, 46:189–194, 2007.

[7] P.C. Hansen and M. Saxild-Hansen. AIR Tools - a MATLAB package of algebraic iterative reconstruction methods. *Journal of Computational and Applied Mathematics*, 236:pp. 2167–2178, 2012.

[8] A.N. Krylov. On the numerical solution of the equation by which the frequency of small oscillations is determined in technical problems. *Izv. Akad. Nauk SSSR Ser. Fiz.-Mat.*, 4, 1931.

[9] Y. Saad and M.H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7:856–869, 1986.

[10] Z. Sun and W.C. Karl. A non-rotational approach to computed tomography. *Poster available online: http://www.censsis.neu.edu/RICC/2010/posters/pdf/HSD_p7.pdf*, 2010.

[11] H.A. van der Vorst and C. Vuik. The superlinear convergence of behaviour of GMRES. *J. Comput. Appl. Mathematics*, 48:327–341, 1993.

[12] J. Wu and Z. Ma. A regularised GMRES method for inverse blackbody radiation problem. 2012.

[13] J.F. Yin and K. Hayami. Preconditioned GMRES methods with incomplete Givens orthogonalization method for large sparse least squares problems. *J. of Comput. and Appl. Mathematics*, 226:177–186, 2009.