Numerical analysis of Krylov multigrid methods for stationary advection-diffusion equation

Nikolay M. Evstigneev

Federal Research Center "Computer Science and Control", Institute for System Analysis, Russian Academy of Science, 117312, Moscow, pr. 60-letiya Oktyabrya, 9, Russia.
E-mail: evstigneevn@yandex.ru

Abstract. This paper outlines the problem of applying multigrid methods for the stationary scalar advection-diffusion equation for the given advection field. We assume that all functions are smooth enough to be represented by discretization methods. Two main discretization are considered: finite difference method for Dirichlet boundary conditions and pseudo-spectral method for the periodic boundary. We test the following smoothers for multigrid methods: Jacobi smoother, Gauss-Seidel smoother, Krylov subspace smoother (GMRES method with and without preconditioners). The analysis is performed in the space formed by the cross product of discretization parameters, diffusion coefficient values, multigrid levels and smoothers.

We demonstrate that the most efficient strategy depends on parameter value and given velocity field. Best variants include Gauss-Seidel smoothers which is optimal for advection-dominated problem while multigrid method is used as a preconditioner for a Krylov method. Such methods can be used for spectral or pseudo-spectral methods where explicit dense matrix storage is impossible.

1. Introduction

This paper outlines the problem of solving numerically stationary advection-diffusion equation:

\[(a, \nabla)u - \mu \Delta u - f = 0\]  \hspace{1cm} (1)

in a domain \( \Omega \subset \mathbb{R}^2 \) with boundary \( \partial \Omega \). Here \( a \) is a given advection vector field, \( \mu \) is a parameter, \( f \) is a source term. Such equations are usually used as a model for more sophisticated physical phenomena. The equation is also used in the momentum conservation part of the Navier-Stokes equations. So efficient numerical methods for this kind of equations are essential for successfully solution of more complicated problems.

There is an extensive literature exist on the solution of linear systems by multigrid methods obtained from the discretization of (1). This includes classical works where a geometric multigrid method is used as a solver [1, 2] or a preconditioner [3, 4]. Analysis of multigrid methods for the reaction–diffusion and advection–diffusion problems is given in [5, 6]. Cascadic variant of a geometric multigrid solver is discussed in [7]. Comparison of geometric and algebraic multigrid methods is given in [8] and Fourier analysis in [9]. A wavelet multigrid method for nonlinear parabolic equations is suggested in [10]. A variant of a multigrid solver for the graphical processing unit (GPU) computational architecture is presented in [11]. High order multigrid method for 3D advection–diffusion equation is discussed in [12, 15]. There is also substantial
research on the algebraic multigrid method, e.g. [13, 14]. Application of monotone schemes in the framework of the multigrid method can be found in [16]. These type of solution methods can be used in real time flow visualization on modern GPUs, e.g. [17].

However, there is no information of the multigrid application for pseudo–spectral methods. One faces this kind of problem while solving eigenvalue problem by iterative methods or while finding a steady state solution of the Oseen problem in periodic domains. In this case a good solver is needed on each iteration of an eigensolver or a Newton’s method. We address this problem in the paper by using the simplest geometric multigrid method (GMG) either as solver or as a preconditioner. Another question is how efficient one can utilize simplest numerical schemes with the GMG on parallel computational architecture (where simple methods can be easily parallelized).

The paper is laid out as follows: first we give a general description of discretization methods and choices of GMG smoothers. Next we consider the problem for (1) in bounded domain with Dirichlet conditions using finite differences for (1) discretization. Then we consider a periodic domain where a pseudo–spectral method is used for discretization of (1). In each section we perform comparison of results and choose best fitting smoothers.

2. Problems and method of solution

2.1. Problem formulation

We are considering two main problems:

(i) Domain \( \Omega = [-1/2, 1/2] \times [-1/2, 1/2] \), Dirichlet boundary conditions \( u|_{\partial \Omega} = 0 \), two variants of advection velocities \( a_1 = (\sqrt{2}/2, \sqrt{2}/2)^T \) and \( a_2 = (y \cos(x/(r + \varepsilon)) − \sin(y/(r + \varepsilon)), x \cos(y/(r + \varepsilon)) − \sin(x/(r + \varepsilon)))/(r^2 + \varepsilon)^T \) with \( \varepsilon = 1.0 \cdot 10^{-12} \) and \( r = \sqrt{x^2 + y^2} \) are considered, the right hand side \( f = -\exp(-C((x - a)^2 + (y - b)^2)) \) with \( C = 60, a = 0.2, b = 0.7 \).

(ii) Domain \( \Omega = [-\pi, \pi] \times [-\pi, \pi] \), periodic boundary conditions on \( u|_{\partial \Omega} \), two advection velocities: \( a_1 = (\sin(y), \cos(x))^T \) and \( a_2 \) is the same, the right hand side \( f = \sin(x) \cos(y) \).

The first problem is solved using finite difference method with the diffusion term discretized using 4th order compact differences scheme [18] and advection term discretized using QUICK or WENO schemes. This results in the system of linear equations:

\[
(N(a) - \mu C) u = f, \tag{2}
\]

where \( N \) is the advection matrix, \( C \) is the diffusion matrix and \( u \in \mathbb{R}^N \).

The second problem is solved using pseudo–spectral method in the Fourier domain resulting in the following equation:

\[
i \sum_{q=1}^{n/2+1} (k_x - q_x) \hat{a}_x q \hat{u}_{k-q} + i \sum_{q=1}^{n/2+1} (k_y - q_y) \hat{a}_y q \hat{u}_{k-q} = -\mu k^2 \hat{u}_k + \hat{f}_k, \tag{3}
\]

for all \( k \), where \( n \) is the number of Fourier harmonics in each direction and the convolution term is treated in the pseudo–spectral manner with 2/3 antialiasing approach. This discretization can also be written in the form (2) where the matrix \( C \) is diagonal and the operator \( N(a) \) is treated as matrix–vector application. Both approaches lead to the standard linear system:

\[
Au = f, \tag{4}
\]

with \( u \in \mathbb{R}^N \) and non-symmetric linear advection-diffusion discrete operator \( A = (N(a) - \mu C) \in \mathbb{R}^{N \times N} \) is treated in the matrix–free manner. Application of Krylov subspace methods without preconditioning to the system (4) results in either poor or no convergence, especially for small values of \( \mu \).
2.2. Solution methods

We use a standard geometric multigrid method with the recursive V-cycle, see e.g. [19]. We shall denote the full V-cycle application of the multigrid method to the vector $u$ as $\mathcal{P}u$.

Prolongator and restrictor operators are defined by linear interpolation for GMG application in physical space. In the Fourier space we use Fourier-wise restriction and prolongation, namely, we restrict higher wave-numbers and prolongate lower ones. This is done by direct injection into the residual array with respect to the data storage (fftw--type in our case).

On each smoothing step of the multigrid method we use a smoothing operator that is executed iteratively on the residual. We use linear iterative smoothers (denoted $LS$ in general), namely, Jacobi smoother (denoted $JS$) and Gauss-Seidel smoother (denoted $GS$) and nonlinear iterative Krylov smoothers (denoted $KS$ in general), namely, GMRES (denoted $GRS$), BiCGStab (denoted $BS$) and BiCGStab(L) (denoted $BLS$) methods. Each of the Krylov methods can have a preconditioner on its own. For example $GRS(GS)$ denoted a GMRES method used as a smoother with the Gauss-Seidel preconditioner. Note, that Krylov subspace methods can hardly be called smoothers in classical meaning of the term, but they can minimize the residual on a Krylov subspace of linear operators [20].

The solution of the main problem (4) is done in two ways. We use a Multigrid method as a solver iteratively:

$$u^{l+1} = \mathcal{P}u^l, l = 1, 2, 3...$$

with $u^0 = f$ and partly factored linear operator in the smoothing step is exactly $A$.

Next, we use a Krylov subspace method to solve the main problem with the multigrid preconditioning, where a linear operator in the multigrid can be an approximation of $A$ in some sense. For any Krylov subspace method we shall have two operations on each iteration:

$$z = Au^l - f,$$

$$r^l = \mathcal{P}z.$$  (6)

In this case we supply a Krylov subspace method with the function handle to the multigrid method.

3. Results

We limit ourselves to the grid resolution up to $512 \times 512$ which decreases as power of two and diffusion coefficient is limited by $R = 1/\mu \leq 10000$. We perform iterations in both methods until at least one condition is met: residual norm is below tolerance ($1.0 \cdot 10^{-9}$), maximum number of iterations is reached ($500$) or residual is greater than threshold ($10\|Af\|$). In all calculations the GMRES method is used without restart.

3.1. First problem

First, we formulate the GMG method as an iterative solver (5) by using simple iteration. Solutions of the problem for $R = \{1000, 10000\}$ and different velocities are presented in figure 1. We observed that the convergence on the velocity field $a_2$ is more difficult so we shall benchmark all results on this velocity.

First we present results of convergence for linear smoothers. Note that usage of each method as a solver alone without GMG resulted either in divergence of iteration process or in prohibitively many iterations. For each multigrid level we use a single iteration of a smoother. Convergence results for $R = 1000$ and $10000$ are shown in figure 2. There are no convergence problems for lower values of $R$. We may notice that $JS$ behaves poorly, as expected resulting in divergence of the iteration process for high values of $R$. The $GS$ smoother is more efficient resulting in convergence for all values of $R$. However the multigrid method cannot guarantee independence of convergence rate as function of grid spacing and diffusion coefficient.
Figure 1. Solution to the first problem for different values of $R$, velocity $a_1$ on the left, velocity $a_2$ on the right.

Then we test nonlinear smoothers in configurations $KS(LS)$. Results are presented in figure 3. On each smoothing level we perform a single iteration with $KS$ that is preconditioned with a single iteration of $LS$. Both $BS$ and $BSL$ lead to divergence for any linear preconditioner. This is expected since both methods minimize residual in a dual spaces and don’t posses smoothing properties. The combination $GRS(JS)$ convergence for most of the problems. It is interesting to note that those problems that fail to converge with $JS$ do converge with $GRS(JS)$. Stagnation near the final residual can be caused by insufficient GMRES tolerance and can be cured by setting higher tolerance for $GRS$ smoothers. The combination of $GRS(GS)$ is more successful resulting in convergence of all problems with the decrease of cumulative iterations.

Figure 2. Convergence of GMG with linear smoothers for different grids and diffusion coefficient values, $GS$ on the left, $JS$ on the right.

Figure 3. Convergence of GMG with nonlinear smoothers for different grids and diffusion coefficient values, $GRS(GS)$ on the left, $GRS(JS)$ on the right.

Now we test Krylov methods with GMG preconditioner that acts as show in (6). First we test $LS$, convergence results are resented in figure 4. We can see that both methods converge successfully with $GS$ smoother being more efficient. Application of other Krylov methods gives similar results. The scheme with Krylov solver and $KS(LS)$ smoothers only works if $KS := GRS$ and a linear preconditioner is used. It results in stagnation of Krylov solvers, see figure 5, or divergence of iterations for high value of $R$.

Finally we bring together all execution times for successfully converged methods in table 1. We can see that the most efficient methods are Krylov solvers with GMG preconditioners where
Gauss-Seidel smoother is used. We can also observe that simple GMG solvers are more grid independent than Krylov solvers, however convergence wall time is much larger. We can also observe that Krylov methods are better suited for cases with low diffusion coefficient (parameter $W_T R = 10000$) giving about 20% increase of wall time.

Table 1. Total wall time ($WT$) and time ratios for all methods that converged to the solution for given grids and given $R$ values for the first problem.

|               | $max\ (WT)$, sec | $WT_{256}$ | $WT_{128}$ | $WT_{512}$ | $WT_{256, R=10000}$ | $WT_{128, R=10000}$ | $WT_{512, R=10000}$ |
|---------------|------------------|------------|------------|------------|---------------------|---------------------|---------------------|
| GMG GS        | 163.5            | 7.61       | 8.21       | 1.41       |
| GMG GRS(GS)   | 305.6            | 6.77       | 8.22       | 1.28       |
| GMRES JS      | 61.9             | 4.81       | 5.88       | 2.23       |
| GMRES GS      | 12.4             | 4.04       | 6.2        | 1.21       |
| BICG GS       | 16.7             | 5.73       | 5.86       | 1.21       |
| BICGL GS      | 17.9             | 5.93       | 6.21       | 1.22       |

3.2. Second problem

Solutions of the second problem for different velocity fields are presented in figure 6. We also observed that the convergence is more difficult for velocity $a_2$, so we are considering all the timings for that velocity field. We observed that pure GMG solver fails for the Fourier method or converges for GMRES smoothers with large number of smoother iterations. So we don’t consider this method here.

Krylov solvers with multigrid preconditioners are formed as follows. First we assemble an operator $\hat{A}$ in physical space that approximates the original operator in the Fourier space. Mapping between two operators is done via Discrete Fourier Transfer, so $\text{DFT}^{-1}(\hat{A}\text{DFT}(x)) \sim Ax$. This operator is assembled using finite difference method analogous to the one used in the
Figure 6. Solutions to the second problem for different values of $R$, velocity $a_1$ on the left, velocity $a_2$ on the right.

First problem but with periodic boundary conditions. This operator is used to construct linear smoothers on GMG levels. The application of $\mathcal{P}$ to the residual in the Krylov solver is done as follows:

$$z = Au^l - f,$$

$$r^l = \text{DFT} \left( \mathcal{P} \left( \text{DFT}^{-1}(z) \right) \right).$$

(7)

Results for Krylov solver with linear preconditioners are presented in figure 7. We can observe that $GS$ smoother is basically more efficient than $JS$ smoother for this problem. However for some intermediate grid sizes $GS$ is less efficient and for the grid $64 \times 64$ and $R = 10000$ the solver failed to converge. For the biggest grid we observed a convergence failure with the given number of iterations for the $JS$ smoother.

Figure 7. Convergence of GMRES with GMG preconditioner and linear smoothers for different grids and diffusion coefficient values, $GS$ on the left, $JS$ on the right.

One can observe an interesting fact, that number of iterations decrease with the grid refinement. We shall consider this problem in the analysis section. Usage of BiCGStab and BiCGStabL gives similar results and graphs are not provided here. Wall times and its ratios for different value of $R$ and grid resolutions are presented in table 2. Again, we can observe anomalous behaviour for smaller grids compared to the bigger ones.

|                | max $(WT)$, sec | $\frac{WT_{R=1000}}{WT_{R=256}}$ | $\frac{WT_{R=128}}{WT_{R=256}}$ | $\frac{WT_{R=512}}{WT_{R=256}}$ |
|----------------|-----------------|-----------------------------------|-----------------------------------|-----------------------------------|
| $GMRES$ $GS$   | 109.3           | 0.98                              | 2.58                              | 0.86                              |
| $GMRES$ $JS$   | 260.4           | 6.24                              | —                                 | —                                 |
| $BICG$ $GS$    | 132.2           | 1.01                              | 3.01                              | 0.98                              |
| $BICG$ $GS$    | 145.0           | 1.12                              | 2.89                              | 1.01                              |
4. Analysis
The solution of the first problem is basically well understood in cited literature and requires no additional analysis. The best solution strategy is the combination of Krylov subspace iterative methods with GMG as a preconditioner where a Gauss-Seidel smoother is used. However we still have dependence on grid parameters and small dependence (about 20% with 10 times change) on the diffusion coefficient.

The second problem needs more attention. We have observed the anomaly in the convergence where a smaller grid suffers worse convergence then the finer one. We consider the spectrum of the problem for $R = 1000$ that is constructed numerically. The results for grid $32 \times 32$ are presented in figure 8 and for grid $64 \times 64$ in figure 9. This combination of diffusion coefficient and grid parameters corresponds to the poorly and well converging grids, see figure 7, first subfigure from the left.

![Figure 8](image1.png)

**Figure 8.** Spectrum of the original matrix $\mathbf{A}$ on the left, spectrum of $\mathbf{P}$ (circles) and $\mathbf{A}^{-1}$ in the middle and spectrum of $\mathbf{PA}$ on the right for grid $32 \times 32$, velocity $a_2$ and $R = 1000$.

![Figure 9](image2.png)

**Figure 9.** Spectrum of the original matrix $\mathbf{A}$ on the left, spectrum of $\mathbf{P}$ (circles) and $\mathbf{A}^{-1}$ in the middle and spectrum of $\mathbf{PA}$ on the right for grid $64 \times 64$, velocity $a_2$ and $R = 1000$.

Results clearly indicate that the matrix for the first grid has unstable spectrum that is not observed in the preconditioner matrix. Such behaviour is not physical (since the spectrum of the same matrix is in the left half plane for the better resolved problem). As the result, the spectrum of the preconditioned matrix is not well–clustered near the point $(1, 0)$ and is not separated from the imaginary axis. This leads to poor GMRES convergence, see [21]. So the explanation for the observed anomaly is due to the noise in the Fourier under-resolved grid. Application of post filters to the original discrete scheme or increase of resolution cures this problem.

It is also observed that for the last two bigger grids number of iterations decreasing with the increase of resolution. This is due to the fact that the difference between two discrete operators $\hat{\mathbf{A}}$ and $\tilde{\mathbf{A}}$ approximating the problem is reducing. So the spectrum of both matrices clusters closer and resulting preconditioned matrix becomes better conditioned.
5. Conclusion
We present a numerical analysis for stationary advection–diffusion equation in physical and Fourier domains with geometric multigrid as solvers or as preconditioners to the Krylov subspace iterative methods. We conclude that a simple geometric multigrid can be a useful tool in accelerating convergence of linear systems for such operators. The best variants is the usage of Krylov subspace iterative methods with Gauss-Seidel smoother. We suggest the usage of a preconditioner that acts in physical space for the Fourier pseudo–spectral approximation of periodic problems. We also observed that poor resolution of problems with Fourier methods can lead to problems in linear system convergence. Note that these results cannot be directly extrapolated for two or three dimensional vector valued functions unless block Gauss-Seidel smoothers are used.

All source codes for these tests are available at author’s GitHub and can be provided by request.

Acknowledgments
This work was supported by the Russian Foundation for Basic Research (grant No. 18-29-10008 mk).

References
[1] Murli M. Gupta, Jules Kouatchou, and Jun Zhang. A Compact Multigrid Solver for Convection-Diffusion Equations. J. comp. phys., 132, 123–129, (1997).
[2] Irad Yavneh, Cornelis H. Venner, and Achi Brandt. Fast Multigrid Solution of the Advection Problem with Closed Characteristics. SIAM J. Sci. Comput., 19(1), 111–125, (1998).
[3] Ramage A. A multigrid preconditioner for stabilised discretisations of advection-diffusion problems. J. of Comp. Appl. Math. 110, 187–203, (1999).
[4] Jie Shen, Feng Wang, Jinchao Xu. A finite element multigrid preconditioner for Chebyshev–collocation methods. App. Num. Math. 33, 471–477, (2000).
[5] Olshanskii M.A. and Reusken A. On the Convergence of a Multigrid Method for Linear Reaction-Diffusion Problems. Computing 65, 193–202, (2000).
[6] Reusken A. Convergence analysis of a multigrid method for convection-diffusion equations. Numerische Mathematik, 91(2), 323–349, (2002).
[7] Xuejun Xu, Ales Janka, Jean-Antoine Desideri. Cascadic multigrid for the advection-diffusion equation. Research Report RR-4574, INRIA, (2002).
[8] Chin-Tien Wu, Elman H. Analysis and Comparison of Geometric and Algebraic Multigrid for ConvectionDiffusion Equations. SIAM J. Sci. Comp. 28(6), 2208–2228, (2006).
[9] Elman H.C. and Ramage A. Fourier Analysis of Multigrid for a Model Two-Dimensional Convection-Diffusion Equation. BIT Num, Math., 46(2), 283–306, (2006).
[10] Doreen De Leon. A new wavelet multigrid method. J. Comp. Appl. Math., 220, 674 – 685, (2008).
[11] Evstigneev, N.M.: Numerical integration of Poisson’s equation using a graphics processing unit with CUDA-technology. Vychisl. Metody Programmn.,10,2, 268–274, (2009).
[12] Yin Wang and Jun Zhang. Fast and Robust Sixth Order Multigrid Computation for 3D Convection Diffusion Equation. J Comput Appl Math. 234(12), 3496–3506, (2010).
[13] Yvan Notay. Aggregation-Based Algebraic Multigrid for Convection-Diffusion Equations. SIAM J. Sci. Comput.,34(4), A2288–A2316, (2012).
[14] S. Chaabane Khelifi, N. Mechtoua, F. Hulsemann, F. Magoules. A hybrid multigrid method for convection-diffusion problems. J of Comp. Appl. Math. 259, 711–719, (2014).
[15] Faizal Haq, Noor Badshah, Kamal Shah, Muhammad Shahzad. Multigrid method for 2D convection diffusion equation based on fourth order compact scheme. Sci.Int.(Lahore),27(6),4975–4980, (2015).
[16] Bayramov N.R. and Kraus J.K. Multigrid methods for convection–diffusion problems discretized by a monotone scheme. Comp. Meth. in Appl. Mech. and Eng., 317, 723–745, (2017).
[17] Feifei Wan, Yong Yin, Suiyu Zha. 3D Parallel Multigrid Methods for Real-Time Fluid Simulation. 3D Res, 9(8), (2018).
[18] Lele, S.K.: Compact Finite Difference Schemes with Spectral-like Resolution. Journal of Computational Physics, 103, 16–42 (1992).
[19] Olshanskii M.A. Lecture notes on multigrid methods. Moscow: Institute of Numerical Mathematics, RAS 2012.
[20] Matinfara M., Zareamoghaddamb H., Eslamia M., Saeidyc M. GMRES implementations and residual smoothing techniques for solving ill-posed linear systems. Comp. Math. with Appl., 63(1), 1–13, (2012).
[21] Saad Y. Iterative Methods for Sparse Linear Systems. Second edition, SIAM, (2003).