IMA Journal of Numerical Analysis (1988) **8**, 273–279

Fourier Analysis of the SOR Iteration

RANDALL J. LEVEQUE
Department of Mathematics, University of Washingon, Seattle, WA 98195, U.S.A.

LLOYD N. TREFETHEN
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

[Received 28 July 1987]

The SOR iteration for solving linear systems of equations depends upon an overrelaxation factor ω. We show that, for the standard model problem of Poisson's equation on a rectangle, the optimal ω and corresponding convergence rate can be obtained rigorously by Fourier analysis. The trick is to tilt the space–time grid so that the SOR stencil becomes symmetrical. The tilted grid also gives new insight into the relationships between the Gauss–Seidel and Jacobi iterations and between the lexicographic and red–black orderings, and into the modified equation analysis of Garabedian.

1. Introduction

FOURIER analysis has been used for nearly fifty years to test the stability of time-dependent finite-difference formulae—the 'von Neumann method'.[9] More recently, it has also become a standard tool for estimating the convergence rate of multigrid iterations.[3] But the analysis of the more classical iteration known as successive overrelaxation—SOR—has been carried out by other means.[5,6,10,11] The reason is that the behaviour of SOR, unlike the other two problems, is dominated by low-frequency modes that are controlled by boundary conditions. The obvious application of Fourier analysis treats the boundary conditions incorrectly, and leads to an incorrect prediction of the optimal convergence rate.

In this note we show that, if the computational grid is tilted by a certain angle in space and time, then Fourier analysis becomes exact for the standard SOR model problem: the five-point discretization of Poisson's equation on a rectangle with Dirichlet boundary conditions, with the variables taken in the natural (lexicographic) ordering.

The SOR model problem was first analysed by Frankel in 1950.[6] Our approach leads to no quantitative results that Frankel did not have, but makes it clear why the eigenvectors of the SOR iteration have the form they do. This analysis is restricted to the rectangular model problem, so it in no way supplants the much more general theory of matrix iterations developed by Young in the early 1950s.[10,11]

In 1956, Garabedian proposed a new analysis of SOR, which today could be described as an early application of the idea of 'modified equations'.[7] He observed that the SOR iteration is a consistent finite-difference approximation of a time-dependent partial differential equation, so that its rate of convergence should approximate the rate of decay of solutions to that equation. To determine this rate, he introduced a new timelike variable $s = t + \frac{1}{2}x + \frac{1}{2}y$ which reduces the differential equation to a canonical form that can be analysed by Fourier methods. Our tilting of the grid corresponds exactly to Garabedian's introduction of the variable s. Thus, for the SOR model problem, the consideration of a partial differential equation is unnecessary, and indeed the analysis in the discrete domain has the advantage that it is exact rather than approximate. Garabedian's idea, however, provides additional insight. (Neither tilted grids nor modified equations are sufficient to analyse all iteration formulae, such as the SOR iteration for the nine-point Laplacian.[2])

Approximate Fourier analysis of SOR (on the usual grid) has been discussed previously by Kuo[8] and Chan & Elman[4] and probably others. Our tilted grid is also equivalent to the 'data flow times' considered by Adams & Jordan for reasons of parallelizability.[1]

2. Jacobi iteration

Consider the discrete Poisson problem

$$\begin{aligned} \frac{1}{h^2}(u_{j-1,k}+u_{j+1,k}+u_{j,k-1}+u_{j,k+1}-4u_{jk}) = f_{jk} \quad (1 \leq j,k \leq N-1), \\ u_{jk} = F_{jk} \quad (j=0,\, N \text{ or } k=0,\, N), \end{aligned} \tag{1}$$

on the square $[0,\pi]^2$, with $h = \pi/N$. Let $\bar{u}^n_{jk}$ denote the approximation to the exact discrete solution u of (1) at the n th step of an iteration, with corresponding error $v^n_{jk} = \bar{u}^n_{jk} - u_{jk}$. Define also $x_j = jh$ and $y_k = kh$ for $0 \leq j,k \leq N$.

The Jacobi iteration is an example in which Fourier analysis works straightforwardly.[10,11] The errors evolve according to

$$\begin{aligned} v^{n+1}_{jk} &= \tfrac{1}{4}(v^n_{j-1,k}+v^n_{j+1,k}+v^n_{j,k-1}+v^n_{j,k+1}) \quad (1 \leq j,k \leq N-1), \\ v^n_{jk} &= 0 \quad (j=0,\, N \text{ or } k=0,\, N). \end{aligned} \tag{2}$$

Let us consider what solutions of the form $v^n_{jk} = g(\xi,\eta)^n e^{i(\xi x_j+\eta y_k)}$ this iteration admits if we ignore the boundary conditions. We obtain immediately

$$g(\xi,\eta) = \tfrac{1}{4}(e^{-i\xi h}+e^{i\xi h}+e^{-i\eta h}+e^{i\eta h}) = \tfrac{1}{2}(\cos\xi h + \cos\eta h). \tag{3}$$

This is the *amplification-factor function* for the Jacobi iteration. The essential property is that it is an even function of ξ and η:

$$g(\xi,\eta) = g(-\xi,\eta) = g(\xi,-\eta) = g(-\xi,-\eta). \tag{4}$$

If we take as initial data the linear combination

$$\sin\xi x_j \sin\eta y_k = -\tfrac{1}{4}(e^{i(\xi x_j+\eta y_k)} - e^{i(-\xi x_j+\eta y_k)} - e^{i(\xi x_j-\eta y_k)} + e^{i(-\xi x_j-\eta y_k)}), \tag{5}$$

where ξ and η are integers in the range $1 \leq \xi,\eta \leq N-1$, then the homogeneous boundary conditions are satisfied at $n=0$. By (4), it follows that

$$v_{jk}^n = g(\xi, \eta)^n \sin \xi x_j \sin \eta y_k$$

satisfies both the interior formula and the boundary conditions for all $n>0$, and therefore

$$\sin \xi x_j \sin \eta y_k$$

is an eigenvector of the Jacobi iteration with eigenvalue $g(\xi, \eta)$. Since there are $(N-1)^2$ of these functions, and they are linearly independent, they constitute a basis for the set of all interior grid functions v_{jk}. Therefore the asymptotic convergence factor for the Jacobi iteration is exactly

$$\rho^{\text{Jacobi}} = \max_{1 \leq \xi, \eta \leq N-1} |\tfrac{1}{2}(\cos \xi h + \cos \eta h)|.$$

The maximum is attained with $\xi, \eta = \pm 1$ or $\xi, \eta = \pm(N-1)$:

$$\rho^{\text{Jacobi}} = \cos h \approx 1 - \tfrac{1}{2}h^2. \tag{6}$$

3. SOR and Gauss–Seidel iterations

If we attempt the same analysis for the SOR iteration, i.e.

$$v_{jk}^{n+1} = (1-\omega)v_{jk}^n + \tfrac{1}{4}\omega(v_{j-1,k}^{n+1} + v_{j+1,k}^n + v_{j,k-1}^{n+1} + v_{j,k+1}^n), \tag{7}$$

the result is

$$g(\xi, \eta) = (1-\omega) + \tfrac{1}{4}\omega g(\xi, \eta)(e^{-i\xi h} + e^{-i\eta h}) + \tfrac{1}{4}\omega(e^{i\xi h} + e^{i\eta h}),$$

that is,

$$g(\xi, \eta) = \frac{1 - \omega + \tfrac{1}{4}\omega(e^{i\xi h} + e^{i\eta h})}{1 - \tfrac{1}{4}\omega(e^{-i\xi h} + e^{-i\eta h})}. \tag{8}$$

Now, (4) no longer holds. Therefore $g(\xi, \eta)$ does not give us the eigenvalues of the SOR iteration. If we find ξ and η to maximize $|g(\xi, \eta)|$, there is no reason to expect the resulting number to describe the convergence of SOR. As it turns out, this approach produces the correct optimal ω, to leading order in h, but a convergence rate that is too pessimistic by a factor of four.[4]

Figure 1 shows how the situation can be rescued. Think of the SOR iteration as inhabiting a regular grid in two space and one time dimensions (j, k, n). Its stencil connects six points in an asymmetrical pattern, or four points in the one-dimensional case portrayed in the figure. Because of this asymmetry, (4) does not hold. But if we introduce the new 'time' index

$$\nu = 2n + j + k, \tag{9}$$

the stencil becomes symmetrical. Let us look for solutions to (7) of the form

$$v_{jk}^\nu = g(\xi, \eta)^\nu e^{i(\xi x_j + \eta y_k)}. \tag{10}$$

In the j, k, ν variables, (7) becomes

$$v_{jk}^{\nu+2} = (1-\omega)v_{jk}^\nu + \tfrac{1}{4}\omega(v_{j-1,k}^{\nu+1} + v_{j+1,k}^{\nu+1} + v_{j,k-1}^{\nu+1} + v_{j,k+1}^{\nu+1}), \tag{11}$$

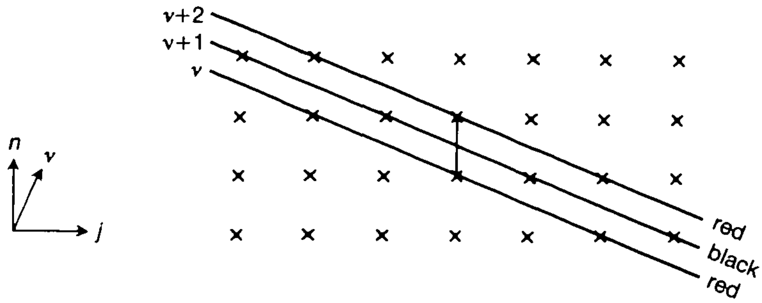

FIG. 1. The SOR stencil superimposed on a space–time grid (one space dimension). Introduction of the 'tilted' time index ν makes the stencil symmetric, so that Fourier analysis can be applied. The red–black labels are explained in Section 4.

and so a suitable value for $g(\xi, \eta)$ is either root of the quadratic equation

$$g(\xi, \eta)^2 = (1-\omega) + \tfrac{1}{4}\omega(\cos \xi h + \cos \eta h) g(\xi, \eta). \tag{12}$$

For each ξ and η, we now have a pair of amplification factors $g_\pm(\xi, \eta)$, and they satisfy the symmetry condition (4). Therefore, for any integers ξ and η in the range $1 \le \xi, \eta \le N-1$, the functions

$$g(\xi, \eta)^\nu \sin \xi x_j \sin \eta y_k$$

are eigenmodes of the SOR iteration in the ν direction. To speak in terms of eigenvectors, we note that SOR is a two-step formula with respect to ν, but it can be recast as a one-step iteration

$$(v^{\nu-2}, v^{\nu-1}) \mapsto (v^\nu, v^{\nu+1})$$

with eigenvectors

$$(\sin \xi x_j \sin \eta y_k, g(\xi, \eta) \sin \xi x_j \sin \eta y_k)$$

and eigenvalues $g(\xi, \eta)^2$.

It takes two steps in ν to advance one step in n. We conclude that the asymptotic convergence factor for SOR is exactly

$$\rho_\omega^{\mathrm{SOR}} = \max_{1 \le \xi, \eta \le N-1} \max_{+,-} |g_\pm(\xi, \eta)|^2. \tag{13}$$

In the original (j, k, n) coordinates, the eigenmodes become

$$v_{jk}^n = g(\xi, \eta)^{2n+j+k} \sin \xi x_j \sin \eta y_k,$$

and the corresponding eigenvectors are

$$g(\xi, \eta)^{j+k} \sin \xi x_j \sin \eta y_k. \tag{14}$$

This matches the results of Frankel and others derived by different means.

All that remains is algebra. For any $\xi, \eta \in \{1, \ldots, N-1\}$, the solutions to (12) are

$$g_\pm(\xi, \eta) = \alpha \pm [\alpha^2 - (\omega - 1)]^{\frac{1}{2}}, \qquad \alpha = \tfrac{1}{4}\omega(\cos \xi h + \cos \eta h), \tag{15}$$

and the larger in magnitude of these two numbers has magnitude

$$\max_{+,-} |g_\pm(\xi,\eta)| = \begin{cases} |\alpha| + [\alpha^2 - (\omega-1)]^{\frac{1}{2}} & \text{if } \alpha^2 \geq \omega - 1, \\ (\omega-1)^{\frac{1}{2}} & \text{if } \alpha^2 \leq \omega - 1, \end{cases} \tag{16}$$

since $\omega \geq 1$. For fixed ω, this quantity can evidently be maximized with respect to ξ and η by taking $\xi = \eta = 1$ (among other values) and $\alpha = \frac{1}{2}\omega \cos h$.

The Gauss–Seidel iteration corresponds to $\omega = 1$. In this case, (16) becomes $2|\alpha| = \cos h$, so by (13) we have

$$\rho^{GS} = \cos^2 h \approx 1 - h^2. \tag{17}$$

To find the optimal overrelaxation factor for SOR, we examine the dependence of (16) on ω with $\xi = \eta = 1$. It is readily verified that $\omega > 2$ leads to $\rho_\omega^{SOR} > 1$, so this is out of the running, and we can assume $1 \leq \omega \leq 2$. In this range, the second line of (16) obviously increases with ω, and differentiation confirms that the first line decreases with ω. Therefore the optimal ω is the crossover value $\omega - 1 = \alpha^2 = (\frac{1}{2}\omega \cos h)^2$, which reduces after a little work to

$$\omega_{opt} = 2/(1 + \sin h) \approx 2 - 2h. \tag{18}$$

By (13) and (16), the corresponding convergence rate is

$$\rho_{opt}^{SOR} = \omega_{opt} - 1 = (1 - \sin h)/(1 + \sin h) \approx 1 - 2h. \tag{19}$$

4. Relating various methods

The change to tilted coordinates has the additional advantage of clarifying the relationships between convergence rates of different iterative methods. For example, it is well known that the Gauss–Seidel iteration is twice as fast as Jacobi, as is confirmed by comparing (6) and (17). The tilted coordinates provide a simple explanation of why this is so. Gauss–Seidel corresponds to the case $\omega = 1$ of (11), and this is precisely the Jacobi iteration with respect to ν. The factor 2 arises because it takes two steps in ν to advance one step in n.

As another example, consider the SOR iteration with the grid points taken in the red–black or checkerboard ordering. This means that the v_{jk} with $j + k$ even (red points) are processed before the v_{jk} with $j + k$ odd (black points), and the iteration takes the form

$$v_{jk}^{n+1} = (1-\omega)v_{jk}^n + \tfrac{1}{4}\omega(v_{j-1,k}^n + v_{j+1,k}^n + v_{j,k-1}^n + v_{j,k+1}^n) \quad \text{for } j + k \text{ even},$$
$$v_{jk}^{n+1} = (1-\omega)v_{jk}^n + \tfrac{1}{4}\omega(v_{j-1,k}^{n+1} + v_{j+1,k}^{n+1} + v_{j,k-1}^{n+1} + v_{j,k+1}^{n+1}) \quad \text{for } j + k \text{ odd}.$$

For each ω, this method has the same convergence rate as the iteration (7) in the natural ordering. Young proved this algebraically by determining the eigenvalues and eigenvectors of the associated iteration matrices.[11] Again, the change to tilted coordinates gives a more intuitive explanation, as illustrated in Fig. 1 for the case of one space dimension. At step ν, we are computing only the red points, and at step $\nu+1$ only the black points. Recasting SOR as a one-step iteration

$$(v^{\nu-2}, v^{\nu-1}) \mapsto (v^{\nu}, v^{\nu+1}),$$

as in the last section, we obtain simply the red–black ordering. Thus Fig. 1 can be viewed as depicting an SOR iteration either in (j, n) coordinates with the natural ordering, or in (j, ν) coordinates with the red–black ordering. Hence these two orderings must have the same asymptotic convergence rate.

The conclusions above depend on the fact that the convergence rate is independent of the particular initial data used, depending only on the eigenvectors. Note that we can switch back and forth between arbitrary data at fixed n or at fixed ν, by taking partial iterations over a triangular portion of the grid. In fact, writing out these partial iterations algebraically gives a similarity transformation relating the iteration matrices.

In this paper, we have considered just the five-point Poisson model problem, and presented an easy way to obtain classical results with, we hope, additional insight. The tilted grid may also prove useful in obtaining new results. It has already been applied to settle a conjecture of Adams & Jordan regarding the equivalence of certain orderings for the nine-point Laplacian.[1] These results are reported in Ref. 2, although as mentioned above, the nine-point formula involves additional complexities.

Acknowledgements

We are grateful to Loyce Adams, Tony Chan, Howard Elman, Jay Kuo and David Young for their advice and encouragement.

The first author was supported by National Science Foundation Grant DMS-8601363 and by the Air Force Office of Scientific Research under AFOSR Contract 85-0189 while in residence at the Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA 23665. The second author was supported by National Science Foundation Grant DMS-8504703, by an IBM Faculty Development Award and U.S. Army Research Office grant DAAL 03-86-K-0171.

REFERENCES

1. ADAMS, L. M., & JORDAN, H. F. 1986 Is SOR color-blind? *SIAM J. sci. stat. Comp.* **7**, 490–506.
2. ADAMS, L. M., LEVEQUE, R. L., & YOUNG, D. M. 1988 Analysis of the SOR iteration for the 9-point Laplacian. *SIAM J. numer. Anal.* (to appear).
3. BRANDT, A. 1977 Multi-level adaptive solutions to boundary-value problems. *Math. Comp.* **31**, 333–390.
4. CHAN, T. F., & ELMAN, H. C. 1987 Fourier analysis of iterative methods for elliptic problems, Univ. of Maryland Dept. of Comp. Sci. Rep. 1763.
5. FORSYTHE, G. E., & WASOW, W. R. 1960 *Finite-Difference Methods for Partial Differential Equations*. Wiley.
6. FRANKEL, S. 1950 Convergence rates of iterative treatments of partial differential equations. *Math. Comp.* **4**, 65–75.

7. GARABEDIAN, P. 1956 Estimation of the relaxation factor for small mesh size. *Math. Comp.* **10,** 183–185.
8. KUO, C.-C. J., LEVY, B. C., & MUSICUS, B. R. 1987 A local relaxation method for solving elliptic PDEs on mesh-connected arrays. *SIAM J. sci. stat. Comp.* **8,** 550–573.
9. RICHTYMER, R. D., & MORTON, K. W. 1967 *Difference Methods for Initial-Value Problems,* 2nd edn. Wiley-Interscience.
10. VARGA, R. S. 1962 *Matrix Iterative Analysis.* Prentice-Hall.
11. YOUNG, D. 1971 *Iterative Solution of Large Linear Systems.* Academic Press.