

The Convergence of Accelerated Overrelaxation Iterations

By A. J. Hughes Hallett

Abstract. Accelerated Overrelaxation Iterations extrapolate the standard Successive Overrelaxation Iterations. This paper provides conditions for the convergence of the Accelerated Overrelaxation process; and a conditionally optimal version is derived for solving an arbitrary real linear equation system.

1. Introduction. Hadjidimos [2] introduced the Accelerated Overrelaxation Iterative method (AOR) for the numerical solution of linear equation systems. It is a two-parameter generalization of the conventional Jacobi, Gauss-Seidel, and Successive Overrelaxation techniques. Convergence results were provided for three cases; where the equation system results in an irreducible and weakly diagonally dominant matrix, or an L -matrix, or a consistently ordered matrix [2]. The purpose of this paper is to extend that convergence analysis to cover *any* real-valued equation system. Some attention is also given to the question of designing an appropriate computational form of the AOR method.

2. AOR Calculations. Consider the system of n equations

$$(1) \quad Ay = b,$$

where $A \in R^{n,n}$ is a known real matrix with nonvanishing diagonal elements, and where y and b are respectively n -vectors of unknown and preassigned real variables. Let $A = D - L - U$ such that

$$D_{ij} = \begin{cases} A_{ij} & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases} \quad L_{ij} = \begin{cases} -A_{ij} & \text{if } j < i, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$U_{ij} = \begin{cases} -A_{ij} & \text{if } j > i, \\ 0 & \text{otherwise.} \end{cases}$$

For convenience we impose the normalization $D = I$. Stationary first-order iterative methods for constructing the numerical solution to (1) have the form

$$(2) \quad y^{(s+1)} = Gy^{(s)} + c, \quad s = 0, 1, 2, \dots,$$

given an arbitrary starting vector $y^{(0)}$. The AOR method is defined by

$$(3) \quad G = (I - \alpha L)^{-1}[(\gamma - \alpha)L + \gamma U + (1 - \gamma)I] = G_{\alpha, \gamma}$$

Received May 2, 1983; revised April 17, 1985 and September 17, 1985.
 1980 *Mathematics Subject Classification.* Primary 65F10.

and $c = \gamma(I - \alpha L)^{-1}b$, where α and γ are scalars. The AOR method is therefore a two-parameter generalization of the most widely used one-parameter first-order iterations. First-order methods are based on the splitting $A = P - Q$, where P is nonsingular, and they are completely consistent with (1) when $G = P^{-1}Q$ and $c = P^{-1}b$ define the iteration matrix and forcing function [9, pp. 64–70]. The AOR iterations contain the following standard methods as special cases,

$G_{0,1}$: the Jacobi method, with $P = I$.

$G_{1,1}$: the Gauss-Seidel method, with $P = (I - L)$.

$G_{0,\gamma}$: the Jacobi overrelaxation method (JOR), with $P = I/\gamma$.

$G_{\alpha,\alpha}$: the successive overrelaxation method (SOR), with $P = (I - \alpha L)/\alpha$.

But the unrestricted AOR method, $G_{\alpha,\gamma}$, is in fact just a one-parameter extrapolation of an SOR iteration:

LEMMA 1. Equation (2), with G defined by (3) for $(\alpha, \gamma) \neq 0$, is identical to

$$(4) \quad y^{(s+1)} = \beta(G_{\alpha,\alpha}y^{(s)} + \alpha c) + (1 - \beta)y^{(s)},$$

where $G_{\alpha,\alpha} = (I - \alpha L)^{-1}(\alpha U + (1 - \alpha)I)$ is the SOR iteration matrix, and $\beta = \gamma/\alpha$.

Proof. Since $L^n = 0$, insert $(I - \alpha L)^{-1}L = \alpha^{-1}\{(I - \alpha L)^{-1} - I\}$ into (3) to establish $G_{\alpha,\gamma} = \beta G_{\alpha,\alpha} + (1 - \beta)I$. Hence (2) and (3) can be rewritten as (4).

Comments: (a) Lemma 1 is important because it identifies the best computational form for the AOR iteration method. In order to take maximum advantage of the sparseness of (1), AOR iterations should be computed in two parts: first the SOR iterations

$$(5) \quad y_j^{(s+1/2)} = \alpha \left\{ \sum_{k=1}^{j-1} L_{jk}y_k^{(s+1/2)} + \sum_{k=j+1}^n U_{jk}y_k^{(s)} + b_j \right\} + (1 - \alpha)y_j^{(s)}$$

for $j = 1, \dots, n$, and then the extrapolation part

$$(6) \quad y^{(s+1)} = \beta y^{(s+1/2)} + (1 - \beta)y^{(s)}.$$

(b) The AOR method is easy to implement. As (5) and (6) show, the sole modifications to any existing SOR computer program are the inclusion of one extra operation (i.e., (6)) and the storage of one extra n -vector. The storage requirements are therefore the same as for JOR methods.

(c) The splitting of A which generates $G_{\alpha,\gamma}$ is given by $P = (I - \alpha L)/\gamma$.

(d) AOR iterations, written as in (4), were called ESOR in [2]. In [5] they were independently proposed as the extrapolated SOR iterations (5) and (6), and their performance studied under the name of Fast Gauss-Seidel (FGS).

3. Convergence Analysis. To establish convergence conditions for AOR iterations we use the following well-known result for one-parameter extrapolations:

LEMMA 2. The JOR iteration $y^{(s)} = \gamma(By^{(s)} + b) + (1 - \gamma)y^{(s)}$ is convergent for some $\gamma > 0$ if and only if $a_j < 1$, $j = 1, \dots, n$, where $B = I - A$ has eigenvalues $\lambda_j = a_j + ib_j$ and $i = \sqrt{-1}$.

Proof. [4, p. 299], or [8, p. 78].

COROLLARY. The JOR iteration of Lemma 2 obviously also converges for some $\gamma < 0$ if $a_j > 1$, all j . But it is divergent if $a_j \leq 1 < a_k$ for $j \neq k$ [8, pp. 73–78].

THEOREM 1. *The AOR iteration, (2) and (3), is convergent for some $(\alpha, \gamma) > 0$ if $a_j < 1, j = 1, \dots, n$, where $B = I - A$ has eigenvalues $\lambda_j = a_j + ib_j$ and $i = \sqrt{-1}$.*

Proof.

$$(7) \quad G_{\alpha, \gamma} = \beta(I + \alpha L + \dots + \alpha^{n-1}L^{n-1})(\alpha U + (1 - \alpha)I) + (1 - \beta)I \\ = \gamma B + (1 - \gamma)I + \gamma \alpha L(B - I) + \gamma O(\alpha^2) \quad \alpha \rightarrow 0,$$

where $O(\alpha^2)$ contains all terms of order α^2 and higher such that $\lim_{\alpha \rightarrow 0} O(\alpha^2)/\alpha = 0$. But $G_{0, \gamma} = \gamma B + (1 - \gamma)I$ has eigenvalues of modulus

$$|\xi_j| = \left[(\gamma(a_j - 1) + 1)^2 + \gamma^2 b_j^2 \right]^{1/2}.$$

This implies $\partial|\xi_j|/\partial\gamma < 0$, both at $\gamma = 0$ and over the interval $\gamma \in [0, (1 - a_j)/r_j^2]$, where $r_j^2 = (a_j - 1)^2 + b_j^2$, if and only if $a_j < 1$ for $j = 1, \dots, n$. Now let $G_{\alpha, \gamma}$ have eigenvalues $\mu_j, j = 1, \dots, n$. By (7), $\lim_{\alpha \rightarrow 0} |\mu_j| = |\xi_j|$. Hence by the above, or by Lemma 2 if $a_j < 1$ all j , some small $\gamma > 0$ exists such that $\{|\mu_j|, |\xi_j|\} < 1$.

COROLLARY. *The SOR iteration converges for some $\alpha > 0$ if $a_j < 1$ all j .*

Proof. Insert $\gamma = \alpha$ in (7) and repeat the argument in the proof of Theorem 1.

Two questions now arise: (i) For what values of α and γ is the AOR iteration convergent? and (ii) What are the optimal values of α and γ ? Unfortunately, only necessary conditions can be placed on α and γ values to imply convergence. The reason for this is as follows. Since $(I - \alpha L)$ is nonsingular, $|G_{\alpha, \gamma} - \mu_j I| = 0$ can be written as

$$(8) \quad |\beta \alpha (U + L - I) + \alpha(\mu_j - 1)L - (\mu_j - 1)I| = 0.$$

This implies $G_{\alpha, \gamma}$ has a characteristic equation which is a polynomial of degree n in both α and μ_j . An explicit solution for μ_j in terms of α is therefore out of the question. In that case, it will not generally be possible to determine the values of α (and hence γ) such that $|\mu_j| < 1$ or that $\max_j |\mu_j|$ is minimized. But necessary conditions are:

THEOREM 2. *AOR iterations are convergent only if $0 < \alpha, \gamma < 2$, when $\alpha, \gamma \neq 0$.*

Proof. Equation (8) indicates that $G_{\alpha, \gamma}$ and $I + \alpha(\mu_j - 1)L - \beta \alpha A$ share eigenvalues. But the latter matrix has trace equal to $n(1 - \alpha\beta) = \sum \mu_j$. Hence, $|\mu_j| < 1$ only if $0 < \gamma < 2$. Secondly, $|G_{\alpha, \gamma} - (1 - \beta)I| = \prod_j (\mu_j - (1 - \beta))$, which implies

$$(9) \quad \prod_j (\mu_j - (1 - \beta)) = |\beta \alpha U + \beta(1 - \alpha)I| = \beta^n (1 - \alpha)^n.$$

Now, combining (9) with the fact that $|\mu_j| < 1$ only if $\beta - 2 < \mu_j - (1 - \beta) < \beta$ yields the desired result that $|\mu_j| < 1$, all j , only if $0 < \alpha < 2$ and $|\gamma| < 2$.

4. Conditional Optimization. There are three computational forms of the AOR iteration which could be used in practice. The form proposed in [2] is

$$(10) \quad y_j^{(s+1)} = \alpha \sum_{k=1}^{j-1} L_{jk} Y_k^{(s+1)} + (\gamma - \alpha) \sum_{k=1}^{j-1} L_{jk} Y_k^{(s)} \\ + \gamma \sum_{k=j+1}^n U_{jk} Y_k^{(s)} + (1 - \gamma) Y_j^{(s)} + \gamma b.$$

Rearranging terms yields

$$(11) \quad y_j^{(s+1)} = \gamma \left(\sum_{k=1}^n B_{jk} y_k^{(s)} + b \right) + (1 - \gamma) y_j^{(s)} + \alpha \sum_{k=1}^{j-1} L_{jk} (y_k^{(s+1)} - y_k^{(s)}),$$

since $L + U = B$, which implies a JOR (γ) iteration plus an error correction mechanism in the final term. Finally, AOR iterations are also just SOR iterations (5) extrapolated by (6). Compared to standard Newton techniques, all three versions offer efficient computational forms in that they fully exploit the system’s sparseness without the need to evaluate and invert matrices of partial derivatives. The point here is that the extrapolated SOR version, (5) and (6), is in many ways the easiest to use. It yields some general convergence results, and it also permits a partial optimization of the iteration parameters α and β .

The number of steps to convergence for the iteration (2), such that

$$\max_j \left| (y_j^{(s+1)} - y_j^{(s)}) / y_j^{(s)} \right| < \tau,$$

is approximately $\log \tau / \log \rho(G)$, where $\rho(G) < 1$ and $\rho(\cdot)$ denotes spectral radius. Hence, the optimal values of α and β (or $\gamma = \beta\alpha$) are those which minimize $\rho(G_{\alpha,\gamma})$. However, for the reasons given in the last section, no explicit relationship between $\rho(G_{\alpha,\gamma})$ and α is available. That leaves only the possibility of an optimal β value conditional on particular α values. Various procedures are now available for computing the optimal extrapolation parameter β in terms of the underlying iteration matrix $G_{\alpha,\alpha}$. Theorem 1 in [6], for example, defines a directed “downhill” search for $\beta^* = \min_{\beta} \{ \max_j |\mu_j| \}$, where μ_j are the roots of $G_{\alpha,\gamma}$, in terms of the eigenvalues of $G_{\alpha,\alpha}$ for a given α value. A more complicated search procedure, which can also be applied to complex equations, appears in [3]. Finally, some simple ways of approximating β^* automatically, without any preliminary eigenvalue calculations, are given in [7] and have proved very effective in test problems.

How useful is this conditional optimality for AOR iterations? Single parameter iterative methods (e.g., JOR, SOR, or (4) with $\alpha = 1$) have been widely used for solving economic models to generate routine forecasts and policy simulations, and the results in [5] show that the use of extrapolation parameters leads to substantial reductions in the cost of solving those models (as measured by the number of iterations to convergence). Hence, the attempt to pick good values for either α or β is certainly worthwhile. The pertinent questions now are: would the additional gains from jointly optimizing α and β also be significant, and how much of those total gains could be obtained by using just one of the parameters? The latter question is important because we have exact results for choosing the extrapolation parameter β , but very few results to guide the choice of the SOR parameter α . Consequently, if the larger gains can be made by choice of β (given a suitable range of α values), or if most of the available improvements can be obtained by choice of β alone, then the optimization of β will obviously be worthwhile. Otherwise, conditional optimality has little practical value for AOR iterations.

In [1], various numerical tests are carried out on the five leading econometric models of the UK economy, and the relevant results appear in Table 1. Evidently, the losses from not using either α or β can be substantial and, bar one uninteresting case, the losses caused by failing to use the extrapolation parameter (col. 2, $\beta = 1$)

TABLE 1
Percentage losses from not optimizing α and β jointly

Model	α, β both optimized	α optimized; $\beta = 1$	β optimized; $\alpha = 1$	$\alpha = 1$; $\beta = 1$
Liverpool	45 iterations ^{a)}	18%	4%	∞ (divergent)
NIESR	13 iterations	0%	3%	4%
CUBS	34 iterations	6%	0%	36%
LBS	20 iterations	6%	0%	22%
HMT	17 iterations	11%	4%	16%

a) average number of iterations per solution in a 20-period forecast. The models are respectively those of Liverpool University, the National Institute of Economic and Social Research, City University, London Business School, and the British Treasury. Full references will be found in [1]. The optimization of α and β (separately or jointly) was carried out here by numerical searches.

are several times larger than the losses from failing to use the SOR acceleration parameter (col. 3, $\alpha = 1$). Conditional optimality is important here because nearly all the available improvements can be generated by choice of the extrapolation parameter β , without any SOR accelerations. It may nevertheless still be useful to conduct a limited search for α values which roughly minimize $\rho(G_{\alpha, \alpha})$.

Department of Economics
 University of Newcastle
 Newcastle-upon-Tyne NE1, 7RU, England

1. P. G. FISHER, S. HOLLY & A. J. HUGHES HALLETT, *Efficient Solution Techniques for Dynamic Rational Expectations Models*, Paper presented to the Econometric Society World Congress, Cambridge, Mass.; and Macroeconomic Modelling Bureau Discussion Paper No. 4, University of Warwick, England.
2. A. HADJIDIMOS, "Accelerated overrelaxation method," *Math. Comp.*, v. 32, 1978, pp. 149-157.
3. A. HADJIDIMOS, "The optimal solution of the extrapolation problem of a first order scheme," *Internat. J. Comput. Math.*, v. 15, 1983, pp. 153-168.
4. A. J. HUGHES HALLETT, *Some Extensions and Comparisons in the Theory of Gauss-Seidel Iterative Techniques for Solving Large Equation Systems* (E. G. Charatsis, ed.), Proceedings of the Econometric Society European Meeting, 1979, North-Holland, Amsterdam, 1981.
5. A. J. HUGHES HALLETT, "Alternative techniques for solving systems of nonlinear equations," *J. Comput. Appl. Math.*, v. 8, 1982, pp. 35-48.
6. A. J. HUGHES HALLETT, "Simple and optimal extrapolations for first order iterations," *Internat. J. Comput. Math.*, v. 15, 1984, pp. 309-318.
7. A. J. HUGHES HALLETT, "Techniques which accelerate the convergence of first order iterations automatically," *Linear Algebra Appl.*, v. 68, 1984, pp. 115-130.
8. E. ISAACSON & H. B. KELLER, *Analysis of Numerical Methods*, Wiley, New York, 1966.
9. D. M. YOUNG, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.