MODELLING AND SIMULATION of systems
P. Breedveld et al. (editors)
J.C. Baltzer AG, Scientific Publishing Co. © IMACS, 1989
pp. 307-309

ITERATIVE METHODS FOR THE SOLUTION OF LINEAR SYSTEMS ARISING FROM SINGULARLY PERTURBED SYSTEMS

Paul A. FARRELL
Department of Mathematical Sciences
Kent State University
Kent, OH 44242, U.S.A.

Abstract: We examine a number of variants of the Gauss-Seidel method for the solution of linear systems arising from a singularly perturbed differential equation. We show that, because of the dominant role played by the characteristics in such problems, the rate of convergence depends critically on solving in the flow direction and thus the spectral radius is an inappropriate measure in this case. We introduce the Symmetric Gauss-Seidel method and demonstrate that it performs well for one and two dimensional examples.

Introduction

We consider the linear systems which arise as a result of the finite difference discretization, on a uniform mesh, of linear singularly perturbed differential equations in one and two dimensions. Such problems are theoretical models for the physical phenomena of convection dominated flow. It is well known that, when the singular perturbation parameter (representing the diffusion coefficient) is small, these problems are difficult to solve numerically due to the prescence of sharp boundary and/or interior layers. One approach to solving these problems is to employ a uniform mesh and either an upwind scheme or an exponentially fitted scheme which attempts to model the boundary layer accurately [1][3]. These methods respect the natural direction of flow in the problem, given by the characteristics of the differential equation.

One Dimensional Problems

We first analyse a sample one dimensional problem.

$$-\epsilon u''(x) + p(x)u'(x) + r(x)u(x) = f(x), \quad 0 < x < 1,$$

$$u(0) = u(1) = 0.$$

We assume that $r(x) \geq 0$ and that $r(x)$ and $p(x)$ have no simultaneous zeros. The behaviour of the problem then depends on the zeros, if any, of $p(x)$ in the interval $[0,1]$. If $p(x) > 0 (< 0)$ the flow is to the right (left) and there is an exponential boundary layer at the left (right). We reflect this fact by discretizing using an upwinded scheme in the forward (backward) direction or exponentially fitted scheme. On a uniform grid of size $h = 1/(n+1)$ the upwinded difference scheme is :

$$-b_i u_{i-1} + a_i u_i - c_i u_{i+1} = f_i, \quad 1 \leq i \leq n,$$

$$u_0 = u_n = 0,$$

where the coefficients a_i, b_i, c_i are given by

$$b_i = \frac{\epsilon}{h^2} + \frac{1}{2h}(|p_i| + p_i) \geq 0, \quad c_i = \frac{\epsilon}{h^2} + \frac{1}{2h}(|p_i| - p_i) \geq 0,$$

$$a_i = \frac{2\epsilon}{h^2} + \frac{1}{h}|p_i| + r_i \geq 0,$$

and $p_i = p(x_i), r_i = r(x_i)$ and $f_i = f(x_i)$. The matrix $A \in R^{n \times n}$, of the difference scheme is thus a tridiagonal matrix. It is easily shown that this matrix is an *irreducibly diagonally dominant M-matrix* [5, p23] under the above assumptions, and thus that $A^{-1} \geq 0$, where the inequality is meant component-wise.

Now consider the matrix splitting of A, into a diagonal D, a lower triangular L, and an upper triangular U matrix, $A = D - L - U$. The forward (backward) point Gauss-Seidel methods, which we shall call FGS and BGS respectively, may now be written as

$$M_f u^{k+1} = N_f u^k + f, \text{ where } M_f = D - L, \ N_f = U,$$

$$M_b u^{k+1} = N_b u^k + f, \text{ where } M_f = D - U, \ N_f = L.$$

In either case, the splitting $A = M - N$ is a *regular splitting* of the matrix A and $\rho(M^{-1}N) < 1$ [5, Theorem 3.14]. Thus both iterative methods are convergent with the same *asymptotic convergence rate*, since $\rho(M_f^{-1}N_f) = \rho(M_b^{-1}N_b) = \rho(D^{-1}(L+U))^{1/2}$. However, although the asymptotic convergence rate is a good measure of the ultimate rate of convergence of these methods, Il'in and Kellogg [2] have shown that the initial rate of convergence for the two methods differ considerably. It is well known that, in general, the error, as measured in the L_∞ norm, may increase initially. For this reason, in [2], the initial error reduction was calculated by considering the matrix norm of the iteration matrix, $||M^{-1}N||_\infty$.

<u>$p(x)$ of one sign e.g. $\bar{p} \geq p(x) \geq \tilde{p} > 0$</u>

In this case, it is shown in [2] that for mesh spacings h which satisfy $\epsilon \leq .8\tilde{p}h$ the error reduction for the *forward* Gauss-Seidel (FGS) is given by

$$||M^{-1}N||_\infty \leq \frac{2\epsilon}{\tilde{p}h^2}, \tag{1}$$

and thus if $\epsilon < \tilde{p}h^2/2$ the error reduction is monotonic. The analysis is performed by writing

$$M^{-1}N = (I - D^{-1}L)^{-1}D^{-1}U$$

and then estimating the norms of $D^{-1}L$ and $D^{-1}U$. These are respectively strictly lower and strictly upper triangular matrices, with a non-zero sub(super)-diagonal whose entries are given by

$$\alpha_i = \frac{\delta + (|p_i| + p_i)/2}{2\delta + |p_i| + r_i h}, \quad \beta_i = \frac{\delta + (|p_i| - p_i)/2}{2\delta + |p_i| + r_i h}, \tag{2}$$

where $\delta = \epsilon/h$. Note that (1) implies very rapid convergence if $\epsilon \ll h$. The *backward* Gauss-Seidel (BGS), on the other hand does not exhibit this property. Conversely, if $p(x) \leq 0$, BGS exhibits rapid convergence under appropriate conditions and FGS does not. Thus, if one solves the linear system in the natural direction given by the characteristics, one achieves rapid convergence. We shall see later that similar problems arise if we assume $p(x)$ has one or more zeros.

We shall now consider another iterative method which we shall call Symmetric Gauss-Seidel (SGS). This consists of an application of FGS, followed by an application of BGS given by:

$$M_f u^{k+1/2} = N_f u^k + f, \quad \text{where } M_f = D - L, N_f = U,$$
$$M_b u^{k+1} = N_b u^{k+1/2} + f, \quad \text{where } M_f = D - U, N_f = L.$$

Thus the iteration matrix is $M_b^{-1}N_b M_f^{-1}N_f$. This may be considered as a special case of the Symmetric Successive Over Relaxation Method (SSOR) with relaxation parameter $\omega = 1$. Thus, as remarked in [4], provided we store intermediate results, it requires no more calculations per iteration than the FGS. Note that the FGS sweep is guaranteed to reduce the error, but the BGS might increase it again. We shall show that this is not the case, although the error reduction is only as good as that for FGS.

<u>THEOREM 1</u>: If the mesh spacing h satisfies $\epsilon < .8\tilde{p}h$ the error reduction for the Symmetric Gauss-Seidel (SGS) is given by

$$||M^{-1}N||_\infty \leq \frac{2\epsilon}{\tilde{p}h^2}.$$

<u>PROOF</u>: This is similar to that of [2]. To simplify notation we write $||.||$ for $||.||_\infty$.

$$\begin{aligned} ||M^{-1}N||_\infty &= ||M_b^{-1}N_b M_f^{-1}N_f|| \\ &\leq ||(I - D^{-1}U)^{-1}|| \, ||D^{-1}L|| \, ||M_f^{-1}N_f||. \end{aligned} \tag{3}$$

It is clear that $D^{-1}U \leq D_0^{-1}U_0$ and $D^{-1}L \leq D_0^{-1}L_0$, where $A_0 = D_0 - L_0 - U_0$ is the matrix corresponding to A with $r_i = 0, \forall i$. Using similar arguments to [2], we have

$$\|M_f^{-1}N_f\| \le \frac{2n\delta}{\hat{p}}\frac{2\delta+\hat{p}}{\delta}\frac{\delta}{2\delta+\hat{p}}, \quad (4)$$

$$\|D^{-1}L\| \le \|D_0^{-1}L_0\| = \frac{\delta+\hat{p}}{2\delta+\hat{p}}, \quad (5)$$

$$1-\|D^{-1}U\| \ge 1-\|D_0^{-1}U_0\| = 1-\frac{\delta}{2\delta+\hat{p}} = \frac{\delta+\hat{p}}{2\delta+\hat{p}}. \quad (6)$$

Using (6), it follows that,

$$\|(I-D^{-1}U)^{-1}\| \le 1/(1-\|D^{-1}U\|) \le \frac{2\delta+\hat{p}}{\delta+\hat{p}}. \quad (7)$$

Combining (3), (7), (5) and (4) we get

$$\|M^{-1}N\|_\infty \le \frac{2\delta+\hat{p}}{\delta+\hat{p}}\frac{\delta+\hat{p}}{2\delta+\hat{p}}\frac{2n\delta}{\hat{p}}\frac{2\delta+\hat{p}}{\delta}\frac{\delta}{2\delta+\hat{p}} \le \frac{2n\delta}{\hat{p}}. \quad \square$$

Thus SGS achieves the same rate of convergence as the FGS, which solves in the characteristic direction. It is easily shown that SGS also achieves the same rate of convergence as BGS for the problem with $p(x) \le \hat{p} < 0$. This is illustrated in Figure 1 and Table 1, which are for the problem

$$-\epsilon u''(x) - u'(x) = -2, \quad 0 < x < 1, \quad (8)$$

$$u(0) = -1, \quad u(1) = 1,$$

with $\epsilon = .000001$. It should be noted that this problem has an exact solution $u(x) = 2x - 1$ and thus exhibits no boundary layers. For all the results exhibited, the initial guess is chosen so that the initial error at each point is 1. Figure 1 gives the log of the error after one iteration at each point of a grid with $h = 1/20$. For this problem, as expected,

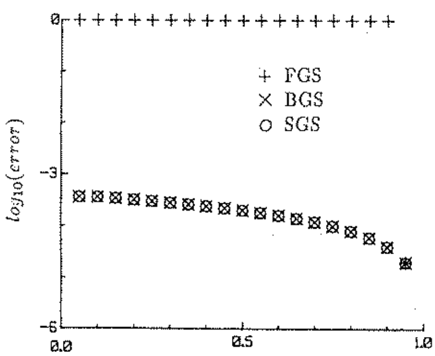

Figure 1: Log of error after first iteration for problem (8)

we get an error of approximately 10^{-4} after the first iteration. The FGS method shows almost no error reduction except at the right-most point where it is 10^{-4}. A similar reduction propagates across the grid at subsequent iterations, requiring a number of iterations of the order of the number of grid-points to reach the left-hand boundary and thus reduce the L_∞ norm significantly. This is illustrated in Table 1.

	Method		
1/h	FGS	BGS	SGS
20	20	2	2
100	101	3	3

Table 1: Problem (8), Number of iterations to achieve error 10^{-5}

<u>$p(x)$ having one or more zeros</u>

There are two interesting cases here. The first is when $p(0) > 0, p(1) < 0$ and $p'(x) < 0$ in $[0,1]$. In this case the zero x^* of $p(x)$ corresponds to a *sink* with the characteristic flow towards x^*. For this case, Il'in and Kellogg [2], used a method which swept to the right while $p_i > 0$ and then sweeps to the left while $p_i < 0$. The second interesting case is when $p(0) < 0, p(1) > 0$ and $p'(x) > 0$ in $[0,1]$. In this case the zero x^* of $p(x)$ corresponds to a *source* and the characteristic flow is directed outward from x^*. In [2], a special block Gauss-Seidel method is proposed, in which the first block is a 2×2 block and the others are 1×1. In both the sink and source cases, it is shown in [2] that the special scheme proposed satisfies a result similar to that for the case where $p(x)$ is of one sign. Both of these special schemes correspond to a reordering of the mesh-points and it is in this context that it is generalized in [2] to the two-dimensional case. It should be noted that neither FGS nor BGS satisfies a monotonicity or rapid convergence result. The SGS, on the other hand, performs well in both cases. This is illustrated in Figure 2 and Table 2 for the sink problem:

$$-\epsilon u''(x) - x/2u'(x) + 3/2u(x) = -1, \quad 0 < x < 1, \quad (9)$$

$$u(0) = -1, \quad u(1) = 1.$$

We can prove a theorem similar to Theorem 1, for this case also.

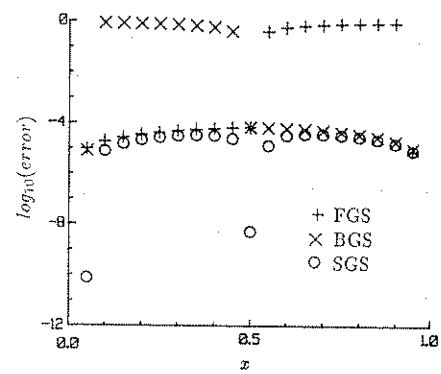

Figure 2: Log of error after first iteration for problem (9)

	Method		
1/h	FGS	BGS	SGS
20	10	10	2
100	50	50	3

Table 2: Problem (9), Number of iterations to achieve error 10^{-5}

<u>THEOREM 2</u>: Let (1) be a sink problem, with the zero, x^*, of $p(x)$ in the interval (x_l, x_{l+1}) and hence, $\hat{p} \ge p(x_i) \ge \mu h$, $\forall i$. Then, for the SGS method, if ϵ satisfies $\epsilon \le \min(.8\hat{p}h, \mu h^3/2)$, the L_∞ norm of the error in iteration $m+1$ satisfies:

$$\|e^{m+1}\| \le \frac{4\epsilon}{\mu h^3}\|e^m\|.$$

<u>PROOF</u>: Consider first the forward sweep. The new error $e^{m+1/2}$ is given by:

$$(I - D^{-1}L)e^{m+1/2} = D^{-1}Ue^m \stackrel{\text{def}}{=} \hat{e},$$

where the entries of $D^{-1}L$ and $D^{-1}U$ are given by (2) and

$$\hat{e} = (\beta_1 e_2^m, \beta_2 e_3^m, \ldots, \beta_{n-1}e_n^m, 0)^T.$$

Divide the system into two parts

$$\begin{pmatrix} B_+ & 0 \\ B_0 & B_- \end{pmatrix}\begin{pmatrix} e_+^{m+1/2} \\ e_-^{m+1/2} \end{pmatrix} = \begin{pmatrix} \hat{e}^+ \\ \hat{e}^- \end{pmatrix}$$

where B_+ is an $l \times l$ matrix and B_- is an $(n-l)\times(n-l)$ matrix. Then

$$B_+e_+^{m+1/2} = \hat{e}^+ \quad (10)$$

$$B_-e_-^{m+1/2} = \hat{e}^- - B_0e_+^{m+1/2} \stackrel{\text{def}}{=} \tilde{e} \quad (11)$$

where $B_0 e_+^{m+1/2} = (-\alpha_{l+1} e_l^{m+1/2}, 0, \ldots, 0)^T$. The analysis for (10) follows as in [2] and Theorem 1, except $\tilde{p}$ is replaced by μh. Thus for $\epsilon \leq .8\tilde{p}h$,

$$\|e_+^{m+1/2}\| \leq \frac{2\epsilon}{\mu h^3}\|e^m\|. \tag{12}$$

For (11), writing $B_- = I - D_- L_-$ and letting D_0 and L_0 be as in Theorem 1 we have

$$\|D_-^{-1}L_-\| \leq \|D_0^{-1}L_0\| = \frac{\delta}{2\delta + \tilde{p}} < 1.$$

Hence

$$\begin{aligned}\|(I - D_-^{-1}L_-)^{-1}\| &\leq 1/(1 - \|D_-^{-1}L_-\|) \leq 1/(1 - \|D_0^{-1}L_0\|) \\ &= \frac{2\delta + \tilde{p}}{\delta + \tilde{p}} \leq 2\,. \end{aligned} \tag{13}$$

Now, using (12) and $\epsilon \leq \mu h^3/2$,

$$\bar{e}_{l+1} = \frac{\delta + |p_{l+1}|}{2\delta + |p_{l+1}| + r_{l+1}h} e_{l+2}^m + \frac{\delta}{2\delta + |p_{l+1}| + r_{l+1}h} e_l^{m+1/2}\,,$$

$$|\bar{e}_{l+1}| \leq \frac{2\delta + |p_{l+1}|}{2\delta + |p_{l+1}| + r_{l+1}h}\|e^m\| \leq \|e^m\|\,, \tag{14}$$

$$|\bar{e}_i| \leq \frac{\delta + |p_i|}{2\delta + |p_i| + r_i h}|e_{i+1}^m| \leq \|e^m\|\,. \tag{15}$$

Thus combining (13), (14), (15) we get

$$\|e_-^{m+1/2}\| \leq 2\|e^m\|\,. \tag{16}$$

Combining (12) and (16) with similar estimates for the backward sweep gives

$$\|e^{m+1}\| \leq \frac{4\epsilon}{\mu h^3}\|e^m\|\,. \quad \square$$

Figure 3 and Table 3 illustrate the results for a source problem:

$$-\epsilon u''(x) + x/2u'(x) + u(x) = 0, \quad 0 < x < 1, \tag{17}$$
$$u(0) = -1, \quad u(1) = 1.$$

In sink and source cases it is clear that for FGS and BGS the error

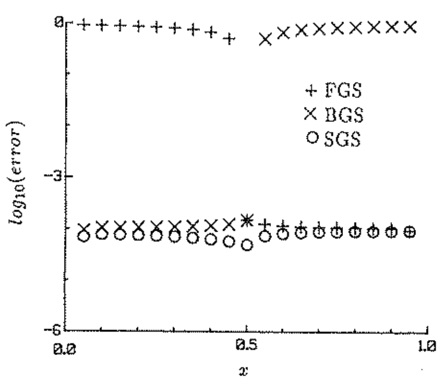

Figure 3: Log of error after first iteration for problem (17)

	Method		
1/h	FGS	BGS	SGS
20	11	11	2
100	52	52	3

Table 3: Problem (17), Number of iterations to achieve error 10^{-5}

is reduced rapidly, where we are solving in the characteristic direction, whereas we reduce the error significantly at only one point when we are solving in the opposite direction. SGS achieves its effect by reducing the error significantly at each point in either the forward or backward sweep.

Two Dimensional Problems

We consider the model two-dimensional convection-diffusion problem on the unit square Ω:

$$-\epsilon\Delta u + p(x,y)u_x + q(x,y)u_y + r(x,y)u = f(x,y),\ (x,y) \in \Omega \tag{18}$$
$$u(x,y) = 0, \ \ (x,y) \in \delta\Omega.$$

We consider five-point upwinded difference schemes for this problem. In [2], it is shown that, if the mesh-points are re-ordered to satisfy an admissibility condition, which is related to the direction of flow at each point, then rapid convergence of the type discussed in the previous section can be achieved by a point Gauss-Seidel method. Asymptotic error results (c.f. [5, Theorem 3.15, p199]) suggest that block methods should achieve better rates of convergence. If the flow is complex, however, a re-ordering to satisfy the admissibility condition is not possible for a block (line) Gauss-Seidel method. This is because the flow at each point of a line is not necessarily in the same direction. In practice as is shown in Table 4, the forward (backward) line Gauss-Seidel FLGS (BLGS) also show the disadvantages of FGS (BGS). We propose a Symmetric Line Gauss-Seidel (SLGS) method which is defined analogously to SGS. The matrix of the difference scheme, A, has the form:

$$\begin{pmatrix} D_1 & U_1 & 0 & \ldots & 0 \\ L_2 & D_2 & U_2 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & L_n & D_n \end{pmatrix}$$

where D_i is a square tri-diagonal matrix of order n, the number of mesh points on each line in the x direction, and L_i and U_i are diagonal matrices of order n. Consider the splitting of A into a block tri-diagonal $D = \text{diag}\,(D_1, D_2, \ldots, D_n)$, a block lower triangular L, and an block upper triangular U matrix, $A = D - L - U$. The SLGS is then defined as :

$$M_f u^{k+1/2} = N_f u^k + f, \quad \text{where } M_f = D - L, N_f = U,$$
$$M_b u^{k+1} = N_b u^{k+1/2} + f, \quad \text{where } M_f = D - U, N_f = L.$$

In practice, under certain restrictions on the coefficients related to stability of the difference scheme, this method achieves rapid convergence results of the type seen for SGS. This is illustrated in Table 4.

			Method		
Problem	Coefficients	1/h	FLGS	BLGS	SLGS
1	$p(x) = 3x - y - 1$	20	2	20	2
	$q(x) = 1$	40	2	40	2
2	$p(x) = 3x - y - 1$	20	23	23	4
	$q(x) = 4x + 2y - 3$	40	40	40	5
3	$p(x) = 3x - y - 1$	20	13	13	3
	$q(x) = -x - 3y - 2$	40	26	26	3

Table 4: Number of iterations to achieve error 10^{-5} for problem (18) with $r(x) = .5$ and $f(x) = 0$.

References

[1] E.P.Doolan, J.J.H.Miller, W.H.A. Schilders, "Uniform Numerical Methods for Problems with Initial and Boundary Layers", Boole Press, Dublin, (1980)

[2] H.Han, V.P.Il'in, R.B.Kellogg, W.Yuan, "Flow Directed Iteration for Convection Dominated Flow", to appear in Proceedings of BAIL V (1988)

[3] R.B.Kellogg, A.Tsan, "Analysis of some Finite Difference Approximations for a Singular Perturbation Problem without Turning points", Math. Comp., 32, pp 1025-1039 (1978)

[4] W.Niethammer, "Relaxation bei Komplexen Matrizen", Math. Zeit., 86, pp 33-40 (1964)

[5] R.S.Varga, "Matrix Iterative Analysis", Prentice-Hall, Englewood Cliffs, NJ, 1962