

Bài 3.3 (20.3) (New) Các phương pháp lặp

Phương pháp khử Gauss và các biến thể của nó trong hai phần cuối cùng là các phương pháp trực tiếp để giải hệ phương trình tuyến tính; đây là những phương pháp đưa ra nghiệm sau một số tính toán có thể được xác định trước.

Ngược lại, trong trường hợp giải gián tiếp hoặc phương pháp lặp chúng ta bắt đầu từ một giá trị xấp xỉ đến nghiệm đúng và, nếu thành công, sẽ có được xấp xỉ tốt hơn và tốt hơn từ một quá trình tính toán lặp đi lặp lại để đạt được độ chính xác cần thiết, số phép tính phụ thuộc vào độ chính xác cần thiết và thay đổi theo từng trường hợp.

Chúng ta áp dụng các phương pháp lặp nếu tốc độ hội tụ là đủ nhanh (nếu ma trận có các phần tử nằm trên đường chéo chính là lớn hơn các phần tử nằm ngoài đường chéo chính, như chúng ta sẽ thấy), chúng ta sẽ phải tính toán ít hơn so với các phương pháp giải trực tiếp.

Chúng ta cũng sử dụng các phương pháp lặp nếu nếu hệ phương trình là lớn và thưa thớt, nghĩa là có rất nhiều hệ số bằng 0, để một trong số đó sẽ lãng phí bộ nhớ để lưu trữ các số 0, ví dụ, trong mỗi phương trình có 9995 số 0 và chỉ có 5 hệ số khác 0 ở hệ có 10.000 phương trình với 10.000 ẩn (ta sẽ nói thêm về điều này trong **Bài 21.4**).

Phương pháp lặp Gauss – Seidel

Đây là phương pháp lặp có tầm quan trọng lớn lao trong các ứng dụng thực tiễn.

Chúng ta sẽ giới thiệu qua ví dụ sau:

Ví dụ 1. Phương pháp lặp Gauss – Seidel

Chúng ta xét hệ pttt

$$\begin{aligned}x_1 - 0.25x_2 - 0.25x_3 &= 50 \\-0.25x_1 + x_2 - 0.25x_4 &= 50 \\-0.25x_1 + x_3 - 0.25x_4 &= 25 \\-0.25x_2 - 0.25x_3 + x_4 &= 25.\end{aligned}$$

(1)

Hệ này xuất hiện trong phép nội suy spline.

Chúng ta viết lại hệ dưới dạng:

$$x_1 = 0.25x_2 + 0.25x_3 + 50$$

$$x_2 = 0.25x_1 + 0.25x_4 + 50$$

$$x_3 = 0.25x_1 + 0.25x_4 + 25$$

$$x_4 = 0.25x_2 + 0.25x_3 + 25.$$

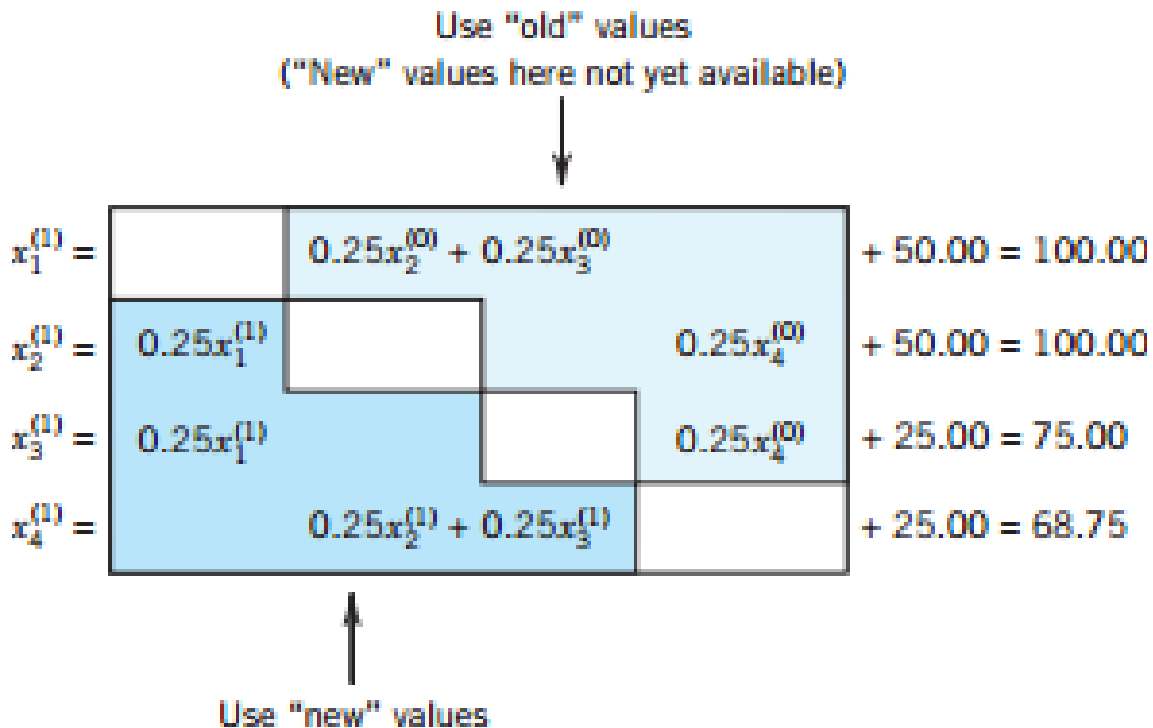
(2)

Hệ pt này được dùng thực hiện phương pháp lặp với giá trị ban đầu:

$$x_1^{(0)} = 50, x_2^{(0)} = 50, x_3^{(0)} = 25, x_4^{(0)} = 25.$$

Và tính từ (2) được giá trị xấp xỉ tốt hơn:

(3)



Hệ phương trình (3) thu được từ (2) bằng cách thay thế ở bên phải xấp xỉ gần nhất cho mỗi ẩn. Trong thực tế, các giá trị tương ứng thay thế các giá trị trước đó ngay khi chúng được tính toán, trong phương trình thứ hai và thứ ba chúng ta sử dụng $x_1^{(1)}$ (chứ không phải $x_1^{(0)}$), trong

phương trình cuối chúng ta sử dụng $x_2^{(1)}$ và $x_3^{(1)}$ (chứ không phải dùng $x_2^{(0)}$ và $x_3^{(0)}$).

Sử dụng nguyên tắc này, chúng ta nhận được kết quả tiếp theo:

$$x_1^{(2)} = 0.25x_2^{(1)} + 0.25x_3^{(1)} + 50.00 = 93.750$$

$$x_2^{(2)} = 0.25x_1^{(2)} + 0.25x_4^{(1)} + 50.00 = 90.625$$

$$x_3^{(2)} = 0.25x_1^{(2)} + 0.25x_4^{(1)} + 25.00 = 65.625$$

$$x_4^{(2)} = 0.25x_2^{(2)} + 0.25x_3^{(2)} + 25.00 = 64.062$$

Và tính toán tiếp

x_1	x_2	x_3	x_4
89.062	88.281	63.281	62.891
87.891	87.695	62.695	62.598
87.598	87.549	62.549	62.524
87.524	87.512	62.512	62.506
87.506	87.503	62.503	62.502

so với nghiệm đúng $x_1 = x_2 = 87,5$,
 $x_3 = x_4 = 62,54$.

Thuật toán lặp Gauss – Seidel được trình bày trong bảng 20.2. Để nhận được thuật toán, chúng ta đưa ra công thức chung cho các số hạng này.

Chúng ta giả thiết rằng, các $a_{jj} = 1$, (điều này là có thể nếu ta xây dựng hệ pt có các hệ số trên đường chéo chính khác 0 và chia mỗi hàng cho phần tử nằm trên đường chéo chính tương ứng). Bây giờ ta viết:

$$(4) \quad A = I + L + U \quad (a_{jj}=1)$$

Ở đây, I là ma trận đơn vị cấp n , còn L và U tương ứng là ma trận tam giác dưới và tam giác trên có các phần tử trên đường

chéo chính bằng 0. Nếu thay thế (4) vào phương trình $Ax=b$ ta có:

$$Ax = (I + L + U)x = b$$

$$(I + L + U)x = b$$

và ta đi đến :

$$(5) \quad x = b - Lx - Ux$$

Nhớ lại từ (3) trong Ví dụ 1, dưới đường chéo chính chúng ta sử dụng các giá trị xấp xỉ “mới” và ở trên đường chéo chính chúng ta sử dụng các giá trị xấp xỉ "cũ", từ (5) chúng ta có được công thức lặp:

$$x^{(m+1)} = b - Lx^{(m+1)} - Ux^{(m)}$$

"New"
"Old"

$$(6) \quad x^{(m+1)} = b - Lx^{(m+1)} - Ux^{(m)} \quad (a_{jj} = 1)$$

"New"
"Old"

$$(I+L)x^{(m+1)} = b - Ux^{(m)}$$

Nhân 2 vế từ bên trái với $(I+L)^{-1}$

Ta nhận được:

$$\begin{aligned} x^{(m+1)} &= (I+L)^{-1}(-U)x^{(m)} + (I+L)^{-1}b \\ &= - (I+L)^{-1}Ux^{(m)} + (I+L)^{-1}b \end{aligned}$$

$$\text{Đặt } C = - (I+L)^{-1}U$$

Công thức lặp Gauss – Seidel:

$$x^{(m+1)} = Cx^{(m)} + (I+L)^{-1}b$$

Ở đây, $x^{(m)} = x_j^{(m)}$ là các giá trị xấp xỉ thứ m, còn $x^{(m+1)} = x_j^{(m+1)}$ là các giá trị xấp xỉ thứ (m+1).

Sau đây là thuật toán lặp Gauss - Seidel

Table 20.2 Gauss–Seidel Iteration

ALGORITHM GAUSS–SEIDEL (\mathbf{A} , \mathbf{b} , $\mathbf{x}^{(0)}$, ϵ , N)

This algorithm computes a solution \mathbf{x} of the system $\mathbf{Ax} = \mathbf{b}$ given an initial approximation $\mathbf{x}^{(0)}$, where $\mathbf{A} = [a_{jk}]$ is an $n \times n$ matrix with $a_{jj} \neq 0$, $j = 1, \dots, n$.

INPUT: \mathbf{A} , \mathbf{b} , initial approximation $\mathbf{x}^{(0)}$, tolerance $\epsilon > 0$, maximum number of iterations N

OUTPUT: Approximate solution $\mathbf{x}^{(m)} = [x_j^{(m)}]$ or failure message that $\mathbf{x}^{(N)}$ does not satisfy the tolerance condition

For $m = 0, \dots, N - 1$, do:

For $j = 1, \dots, n$, do:

$$1 \quad x_j^{(m+1)} = \frac{1}{a_{jj}} \left(b_j - \sum_{k=1}^{j-1} a_{jk} x_k^{(m+1)} - \sum_{k=j+1}^n a_{jk} x_k^{(m)} \right)$$

End

2 If $\max_j |x_j^{(m+1)} - x_j^{(m)}| < \epsilon |x_j^{(m+1)}|$ then OUTPUT $\mathbf{x}^{(m+1)}$. Stop

[Procedure completed successfully]

End

OUTPUT: "No solution satisfying the tolerance condition obtained after N iteration steps." Stop

[Procedure completed unsuccessfully]

End GAUSS–SEIDEL

Bước 1. Đưa ma trận hệ số \mathbf{A} về dạng ma trận có các phần tử nằm trên đường chéo chính bằng 1 và là hệ số lớn nhất trong hàng, nếu có thể.

Bước 2. Tách A thành tổng của 3 ma trận I, L, và U:

$$A = I + L + U$$

Trong đó, I là ma trận đơn vị cấp n

L là ma trận tam giác dưới có các phần tử nằm dưới đường chéo chính.

U là ma trận tam giác trên có các phần tử nằm trên đường chéo trên

Bước 3. Từ pt ban đầu $Ax = b$ ta đi đến phương trình $x = -(I+L)^{-1}Ux + (I+L)^{-1}b$
 $= Cx + d$

Trong đó, $C = -(I+L)^{-1}U$ và $d = (I+L)^{-1}b$

Bước 4. Công thức lặp

Lấy giá trị ban đầu $x^{(0)} = d$

Công thức lặp $x^{(k)} = Cx^{(k-1)} + d$

Bước 5. Điều kiện hội tụ

Chuẩn của ma trận C thỏa $\|C\| < 1$.

Chuẩn của C có thể tính theo 1 trong 3 chuẩn:

- Chuẩn Frobnius:

$$(9) \quad \|C\| = \sqrt{\sum_{j=1}^n \sum_{k=1}^n c_{jk}^2} \quad (\text{Chuẩn Frobnius})$$

- Chuẩn hàng:

$$(11) \quad \|C\| = \max_j \sum_{k=1}^n |c_{jk}| \quad (\text{chuẩn “hàng”})$$

- Chuẩn cột:

$$(10) \quad \|C\| = \max_k \sum_{j=1}^n |c_{jk}| \quad (\text{chuẩn “cột”})$$

Bước 6. Điều kiện dừng

Với $\varepsilon > 0$ bé tùy ý cho trước, quá trình lặp sẽ dừng nếu

$$\|x^{(k)} - x^{(k-1)}\| < \max \{ |x_j^{(k)} - x_j^{(k-1)}|, 1 \leq j \leq n \} < \varepsilon$$

Sự hội tụ và chuẩn của ma trận

Phương pháp lặp để giải hệ pt $Ax=b$ được gọi là hội tụ với giá trị ban đầu $x^{(0)}$, nếu dãy lặp tương ứng $x^{(0)}, x^{(1)}, x^{(2)}, \dots$

hội tụ tới nghiệm của pt đã cho. Sự hội tụ phụ thuộc vào quan hệ giữa $x^{(m)}$ và $x^{(m+1)}$. Để làm rõ mối quan hệ này, ta sử dụng công thức (6). Trước hết, ta có:

$$(I + L)x^{(m+1)} = b - Ux^{(m)}$$

bằng cách nhân phía trái 2 vế với

$$(I + L)^{-1} \text{ ta có:}$$

$$x^{(m+1)} = Cx^{(m)} + (I + L)^{-1} b$$

$$x^{(0)} = (I + L)^{-1} b.$$

$$\text{ở đây } C = -(I + L)^{-1} U.$$

Dãy lặp theo pp Gauss – Seidel luôn hội tụ, nếu và chỉ nếu tất cả các trị riêng

(xem 8.1) của ma trận lặp $C = [c_{jk}]$ có trị tuyệt đối nhỏ hơn 1.

Chú ý: Để nhận được C thỏa mãn yêu cầu, chúng ta cần chia các hàng của A cho a_{jj} , khi đó, các phần tử nằm trên đường chéo chính bằng 1. Nếu bán kính phổ của C (= giá trị lớn nhất của các trị tuyệt đối của các trị riêng) là nhỏ thì tốc độ hội tụ sẽ nhanh.

Điều kiện đủ để dãy lặp hội tụ

Điều kiện đủ để dãy lặp hội tụ là

$$(8) \quad \|C\| < 1$$

ở đây, $\|C\|$ là chuẩn nào đó của ma trận, ví dụ:

$$(9) \quad \|C\| = \sqrt{\sum_{j=1}^n \sum_{k=1}^n c_{jk}^2} \quad (\text{Chuẩn Frobnius})$$

hay, tổng lớn nhất của các $|c_{jk}|$ trong 1 cột của C:

$$(10) \quad \|C\| = \max_k \sum_{j=1}^n |c_{jk}| \quad (\text{chuẩn “cột”})$$

hay, tổng lớn nhất của các $|c_{jk}|$ trong 1 hàng của C:

$$(11) \quad \|C\| = \max_j \sum_{k=1}^n |c_{jk}| \quad (\text{chuẩn “hàng”})$$

Đây là những chuẩn ma trận hay được dùng nhất trong thực tiễn.

Trong hầu hết các trường hợp, việc lựa chọn tiêu chuẩn nào chỉ là vấn đề thuận tiện về mặt tính toán.

Tuy nhiên, ví dụ sau cho thấy rằng đôi khi có tiêu chuẩn thích hợp hơn các tiêu chuẩn khác.

Ví dụ 2. Kiểm tra sự hội tụ của PP lặp Gauss – Seidel

Giải hệ pt

$$2x + y + z = 4$$

$$x + 2y + z = 4$$

$$x + y + 2z = 4$$

Viết lại

$$x = 2 - \frac{1}{2}y - \frac{1}{2}z$$

$$y = 2 - \frac{1}{2}x - \frac{1}{2}z$$

$$z = 2 - \frac{1}{2}x - \frac{1}{2}y$$

Giải

Chia các hàng cho các pt nằm trên đường chéo chính tương ứng, ta được:

$$\begin{bmatrix}
 \mathbf{I} & \mathbf{0} & \mathbf{0} \\
 \mathbf{0} & \mathbf{I} & \mathbf{0} \\
 \mathbf{0} & \mathbf{0} & \mathbf{I}
 \end{bmatrix}
 =
 \begin{bmatrix}
 \mathbf{I} & \mathbf{0} & \mathbf{0} \\
 \mathbf{0} & \mathbf{I} & \mathbf{0} \\
 \mathbf{0} & \mathbf{0} & \mathbf{I}
 \end{bmatrix}
 =
 \mathbf{I} + \mathbf{0} + \mathbf{0} =$$

$$= \mathbf{I} + \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} + \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix}.$$

Khi đó, ta tính được

$$\begin{aligned} \mathbf{C} &= -(\mathbf{I} + \mathbf{L})^{-1} \mathbf{U} = \\ &= - \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ -\frac{1}{4} & -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} = \end{aligned}$$

$$= \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{8} & \frac{3}{8} \end{bmatrix}.$$

Chúng ta tính chuẩn Frobenius của C:

$$\|C\| = \left(\frac{1}{4} + \frac{1}{4} + \frac{1}{16} + \frac{1}{16} + \frac{1}{64} + \frac{9}{64} \right)^{1/2} = \left(\frac{50}{64} \right)^{1/2} = 0.884 < 1$$

và kết luận từ (8) rằng PP lặp Gauss-Seidel hội tụ. Điều thú vị là hai tiêu chuẩn kia sẽ không cho phép kết luận như vậy. Tất nhiên, điều này chỉ ra rằng (8) là đủ để PP lặp hội tụ thay vì cần thiết phải kiểm định theo tất cả các chuẩn.

Sự hội tụ và sai số.

Chúng ta thừa nhận định lý sau, chứng minh của nó là trường hợp riêng của hệ phương trình phi tuyến sẽ được đề cập trong mục 3.1 dưới đây.

Định lý. Nếu đưa được hệ (3.16) về hệ tương đương (3.16') thì hệ chúng có duy nhất nghiệm x^* và $\lim_{k \rightarrow \infty} x^{(k)} = x^*$. Hơn nữa, ta có ước lượng sai số :

$$\left| x_i^* - x_i^{(k)} \right| \leq \frac{q}{1-q} \max_{j \leq n} \left\{ \left| x_j^k - x_j^{k-1} \right| \right\}; \forall i \leq n$$

và ước lượng tiên nghiệm

$$\left| x_i^* - x_i^{(k)} \right| \leq \frac{q^k}{1-q} \max_{j \leq n} \left\{ \left| x_j^1 - x_j^0 \right| \right\}$$

Thông thường ta có thể chọn $x^0 = d$, khi đó coi như ta đã tính xấp xỉ ban đầu với $x^0 = 0$ và $x^0 = d$ là bước tiếp theo.

Phương pháp lặp Jacobi

Phương pháp lặp Gauss-Seidel là một phương pháp hiệu chỉnh liên tiếp bởi vì cho mỗi thành phần ta liên tục thay thế một thành phần của véc tơ xấp xỉ bởi một thành phần xấp xỉ mới vừa tính xong. Một phương pháp lặp được gọi là

một phương pháp hiệu chỉnh đồng thời nếu không có thành phần của $x^{(m)}$ được sử dụng cho đến khi tất cả các thành phần của $x^{(m)}$ được tính toán xong. Phương pháp loại này là Jacobi.

Phương pháp lặp Jacobi tương tự như sự PP lặp Gauss-Seidel, nhưng không giá trị hiệu chỉnh nào được sử dụng cho đến khi tất cả các thành phần được hoàn thành và ta thay thế $x^{(m)}$ bằng $x^{(m+1)}$ một lần, trực tiếp trước khi bắt đầu bước tiếp theo.

Do đó, nếu chúng ta viết hệ $Ax=b$ (với $a_{jj}=1$ như trước) dưới dạng:

$$x = b + (I - A)x$$

Phép lặp Jacobi dưới dạng ma trận có thể viết:

$$\mathbf{x}^{(m+1)} = \mathbf{b} + (\mathbf{I} - \mathbf{A})\mathbf{x}^{(m)}$$

Điều kiện hội tụ: $\|\mathbf{I} - \mathbf{A}\| < 1$.

Điều kiện dừng $\|\mathbf{x}^{(m+1)} - \mathbf{x}^{(m)}\| < \varepsilon$.

PROBLEM SET 20.3

1. Verify the solution in Example 1 of the text.
2. Show that for the system in Example 2 the Jacobi iteration diverges. *Hint.* Use eigenvalues.
3. Verify the claim at the end of Example 2.

4-10 GAUSS-SEIDEL ITERATION

Do 5 steps, starting from $\mathbf{x}_0 = [1 \ 1 \ 1]^T$ and using GS in the computation. *Hint.* Make sure that you solve each equation for the variable that has the largest coefficient (why?). Show the details.

$$4. \quad 4x_1 - x_2 = 21$$

$$-x_1 + 4x_2 - x_3 = -45$$

$$-x_2 + 4x_3 = 33$$

$$5. \quad 10x_1 + x_2 + x_3 = 6$$

$$x_1 + 10x_2 + x_3 = 6$$

$$x_1 + x_2 + 10x_3 = 6$$

$$6. \quad x_2 + 7x_3 = 25.5$$

$$5x_1 + x_2 = 0$$

$$x_1 + 6x_2 + x_3 = -10.5$$

$$7. \quad 5x_1 - 2x_2 = 18$$

$$-2x_1 + 10x_2 - 2x_3 = -60$$

$$-2x_2 + 15x_3 = 128$$

$$8. \quad 3x_1 + 2x_2 + x_3 = 7$$

$$x_1 + 3x_2 + 2x_3 = 4$$

$$2x_1 + x_2 + 3x_3 = 7$$

$$9. \quad 5x_1 + x_2 + 2x_3 = 19$$

$$x_1 + 4x_2 - 2x_3 = -2$$

$$2x_1 + 3x_2 + 8x_3 = 39$$

$$10. \quad 4x_1 \quad \quad + 5x_3 = 12.5$$

$$x_1 + 6x_2 + 2x_3 = 18.5$$

$$8x_1 + 2x_2 + x_3 = -11.5$$

$$8x_1 + 2x_2 + x_3 = -11.5$$

$$x_1 + 6x_2 + 2x_3 = 18.5$$

$$4x_1 \quad \quad + 5x_3 = 12.5$$

11. Apply the Gauss–Seidel iteration (3 steps) to the system in Prob. 5, starting from (a) 0, 0, 0 (b) 10, 10, 10. Compare and comment.
12. In Prob. 5, compute \mathbf{C} (a) if you solve the first equation for x_1 , the second for x_2 , the third for x_3 , proving convergence; (b) if you nonsensically solve the third equation for x_1 , the first for x_2 , the second for x_3 , proving divergence.
13. **CAS Experiment. Gauss–Seidel Iteration.** (a) Write a program for Gauss–Seidel iteration.
- (b) Apply the program $\mathbf{A}(t)\mathbf{x} = \mathbf{b}$, to starting from $[0 \ 0 \ 0]^T$, where

$$\mathbf{A}(t) = \begin{bmatrix} 1 & t & t \\ t & 1 & t \\ t & t & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}.$$

For $t = 0.2, 0.5, 0.8, 0.9$ determine the number of steps to obtain the exact solution to 6S and the corresponding spectral radius of C . Graph the number of steps and the spectral radius as functions of t and comment.

(c) Successive overrelaxation (SOR). Show that by adding and subtracting $\mathbf{x}^{(m)}$ on the right, formula (6) can be written

$$\mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \mathbf{b} - \mathbf{L}\mathbf{x}^{(m+1)} - (\mathbf{U} + \mathbf{I})\mathbf{x}^{(m)} \quad (a_{jj} = 1).$$

Anticipation of further corrections motivates the introduction of an overrelaxation factor $\omega > 1$ to get the SOR formula for Gauss–Seidel

$$(14) \quad \begin{aligned} \mathbf{x}^{(m+1)} = \mathbf{x}^{(m)} + \omega(\mathbf{b} - \mathbf{L}\mathbf{x}^{(m+1)} \\ - (\mathbf{U} + \mathbf{I})\mathbf{x}^{(m)}) \end{aligned} \quad (a_{jj} = 1)$$

intended to give more rapid convergence. A recommended value is $\omega = 2/(1 + \sqrt{1 - \rho})$, where ρ is the spectral radius of \mathbf{C} in (7). Apply SOR to the matrix in (b) for $t = 0.5$ and 0.8 and notice the improvement of convergence. (Spectacular gains are made with larger systems.)

14–17 JACOBI ITERATION

Do 5 steps, starting from $\mathbf{x}_0 = [1 \ 1 \ 1]$. Compare with the Gauss–Seidel iteration. Which of the two seems to converge faster? Show the details of your work.

14. The system in Prob. 4
15. The system in Prob. 9
16. The system in Prob. 10
17. Show convergence in Prob. 16 by verifying that $\mathbf{I} - \mathbf{A}$, where \mathbf{A} is the matrix in Prob. 16 with the rows divided by the corresponding main diagonal entries, has the eigenvalues -0.519589 and $0.259795 \pm 0.246603i$.

18–20**NORMS**

Compute the norms (9), (10), (11) for the following (square) matrices. Comment on the reasons for greater or smaller differences among the three numbers.

18. The matrix in Prob. 10

19. The matrix in Prob. 5

20.
$$\begin{bmatrix} 2k & -k & -k \\ k & -2k & k \\ -k & -k & 2k \end{bmatrix}$$