5.3 Semi-direct and Iterative Methods

Notes

To solve $Ax = b$ when A is difficult to work with:

$$Ax = b \rightarrow Mx = (M-A)x + b$$

Solve iteratively. Guess $x = x_0$ initially and then:

$$\boxed{Mx_{k+1} = (M-A)x_k + b}$$

Choices for M discussed later. If the sequence of vectors converges, i.e. $\lim x_k = x_\infty$, then they approach the true solution of $Ax = b$:

$$Mx_\infty = (M-A)x_\infty + b \rightarrow Ax_\infty = b$$

Error: $\quad Mx - Mx_{k+1} = (M-A)x + b - (M-A)x_k - b$

$$Me_{k+1} = (M-A)e_k \qquad e_j := x - x_j$$

$$e_{k+1} = M^{-1}(M-A)e_k = Be_k$$

Convergence is decided by powers of the matrix B:

$$B = M^{-1}(M-A)$$

The current error is related to $e_0 = x - x_0$ by $e_k = B^k e_0$. Convergence is a question of stability:

$$e_k \to 0 \quad \text{and} \quad x_k \to x \quad \text{if} \quad B^k \to 0$$

<u>5H</u> The powers of B^k approach zero iff $|\lambda_i| < 1$ for each eigenvalue λ_i of B. The rate of convergence is governed by the largest value of $|\lambda_i|$, called the <u>spectral radius</u> of B.

The eigenvalues could be real or complex so for $\lambda = a + ib = re^{i\theta}$, $|\lambda| = \sqrt{a^2+b^2} = r$. Need $r < 1$.

If B has a complete set $x_1, \ldots, x_n$ of eigenvectors,

$$e_0 = c_1\lambda_1 x_1 + \ldots + c_n\lambda_n x_n \rightarrow B^k e_0 = c_1\lambda_1^k x_1 + \ldots + c_n\lambda_n^k x_n$$

This shows that if $B^k \to 0$, $e_k = B^k e_0 \to 0$, which means each $|\lambda_i|$ must be less than 1 (and conversely).

For $B = S\Lambda S^{-1}$, $B^k = S\Lambda^k S^{-1}$. If B has a repeated eigenvalue, try the Jordan Form: $B = PJP^{-1}$ so that $B^k = PJ^kP^{-1}$.

The matrix M

- Jacobi's Method

If all a_{ii} of A are nonzero, take $M = D$, where D is the diagonal matrix with diagonal elements matching the diagonal of A.

$$Dx_{k+1} = (D-A)x_k + b$$

$$\begin{aligned} a_{11}(x_1)_{k+1} &= -a_{12}(x_2)_k - a_{13}(x_3)_k - \dots - a_{1n}(x_n)_k + b_1 \\ &= (-a_{12}x_2 - \dots - a_{1n}x_n)_k + b_1 \end{aligned}$$

$$\begin{aligned} a_{22}(x_2)_{k+1} &= -a_{21}(x_1)_k - a_{23}(x_3)_k - \dots - a_{2n}(x_n)_k + b_2 \\ &= (-a_{21}x_1 - \dots - a_{2n}x_n)_k + b_2 \end{aligned}$$

$$\vdots$$

$$\begin{aligned} a_{nn}(x_n)_{k+1} &= -a_{n1}(x_1)_k - a_{n2}(x_2)_k - \dots - a_{n,n-1}(x_{n-1})_k + b_n \\ &= (-a_{n1}x_1 - \dots - a_{n,n-1}x_{n-1})_k + b_n \end{aligned}$$

This method requires us to store all components of x_k until calculation of x_k is complete. A more natural and efficient idea is to the components of x_k with the components of x_{k+1} as soon as they are found.

This improvement is called Gauss-Seidel (even though Gauss didn't know about and Seidel never recommended it).

- Gauss-Seidel Method

The first component of x_{k+1}, $(x_1)_{k+1}$ is calculated the same way as in Jacobi:

$$a_{11}(x_1)_{k+1} = -a_{12}(x_2)_k - a_{13}(x_3)_k - \dots - a_{1n}(x_n)_k + b_1$$
$$= (-a_{12}x_2 - \dots - a_{1n}x_n)_k + b_1$$

Use this to help calculate $(x_2)_{k+1}$, and so on,

$$a_{22}(x_2)_{k+1} = -a_{21}(x_1)_{k+1} - a_{23}(x_3)_k - \dots - a_{2n}(x_n)_k + b_2$$
$$= -a_{21}(x_1)_{k+1} - (a_{23}x_3 - \dots - a_{2n}x_n)_k + b_2$$

$$a_{33}(x_3)_{k+1} = -a_{31}(x_1)_{k+1} - a_{32}(x_2)_{k+1} - a_{34}(x_4)_k - \dots - a_{3n}(x_n)_k + b_3$$
$$= (-a_{31}x_1 - a_{32}x_2)_{k+1} - (a_{34}x_4 - \dots - a_{3n}x_n)_k + b_3$$
$$\vdots$$

$$a_{nn}(x_n)_{k+1} = (-a_{n1}x_1 - \dots - a_{n,n-1}x_{n-1})_{k+1} + b_n$$

- Successive Over - Relaxation

Example

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \rightarrow M = \begin{bmatrix} 2/\omega & 0 \\ -1 & 2/\omega \end{bmatrix} \quad (\omega > 1)$$

$$M x_{k+1} = (M - A) x_k + b$$

$$B = I - M^{-1} A = \begin{bmatrix} 1-\omega & \frac{1}{2}\omega \\ \frac{1}{2}\omega(1-\omega) & 1-\omega+\frac{1}{4}\omega^2 \end{bmatrix}$$

$$\lambda_1 \lambda_2 = \det B = (1-\omega)^2$$

As one eigenvalue decreases the other increases. In order to minimize $\max |\lambda_i|$,

$$\lambda_1 = \lambda_2 = \omega - 1$$

$$2(\omega-1) = \lambda_1 + \lambda_2 = \text{trace } B = 2(1-\omega) + \frac{1}{4}\omega^2$$

$$\longrightarrow \omega_{opt} = 4(2-\sqrt{3}) \approx 1.07 \longrightarrow \lambda_1 = \lambda_2 = 0.07$$

Power Method for Eigenvalues

5I Ordinary Power Method. Spse A has a single largest eigenvalue. Then if u_0 contains any component $c_n x_n$ of x_n, the sequence $u_k = A^k u_0$ converges (after rescaling) to that eigenvector:

$$\frac{u_k}{c_n \lambda_n^k} = \frac{c_1}{c_n}\left(\frac{\lambda_1}{\lambda_n}\right)^k x_1 + \cdots + \frac{c_{n-1}}{c_n}\left(\frac{\lambda_{n-1}}{\lambda_n}\right)^k x_{n-1} + x_n$$

$$\frac{u_k}{c_n \lambda_n^k} \longrightarrow x_n \quad \text{as} \quad k \longrightarrow \infty$$

The next largest λ gives the convergence factor:

$$r = \frac{|\lambda_{n-1}|}{|\lambda_n|}$$

Exercises

5.3.1

Jacobi iteration for a general 2×2 matrix has

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad M = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

Find the eigenvalues of $B = M^{-1}(M-A)$. If A is symmetric positive definite, show that the iteration converges.

$$B = \begin{bmatrix} 1/a & 0 \\ 0 & 1/d \end{bmatrix} \begin{bmatrix} 0 & -b \\ -c & 0 \end{bmatrix} = \begin{bmatrix} 0 & -b/a \\ -c/d & 0 \end{bmatrix}$$

$$0 = \lambda^2 - \frac{bc}{ad} \longrightarrow \boxed{\lambda^2 = \frac{bc}{ad}}$$

If A is symmetric positive definite,

$$b = c, \quad a > 0, \quad ad > b^2 \geq 0$$

$$\Rightarrow \lambda^2 = \frac{b^2}{ad} > 0 \quad \Rightarrow \boxed{\lambda = \pm\sqrt{\frac{b^2}{ad}}}$$

5.3.2

For Gauss-Seidel the iteration matrices are

$$M = \begin{bmatrix} a & 0 \\ c & d \end{bmatrix} \quad M - A = \begin{bmatrix} 0 & -b \\ 0 & 0 \end{bmatrix}$$

Find the eigenvalues of $B = M^{-1}(M-A)$. Give an example of a matrix A for which Gauss-Seidel will not converge.

- $$B = \frac{1}{ad}\begin{bmatrix} d & 0 \\ -c & a \end{bmatrix}\begin{bmatrix} 0 & -b \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -bd/ac \\ 0 & bc/ad \end{bmatrix}$$

$$0 = -\lambda(bc/ad - \lambda) \longrightarrow \boxed{\lambda = 0,\ bc/ad}$$

- For G-S divergence, need $\lambda = bc/ad \geq 1$

$$a = b = c = d = 1 \iff A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} = B^2 = B^3 = \dots$$

Since $B^k \not\to 0$, G-S diverges for A.

5.3.4 Decide the convergence or divergence of Jacobi and Gauss-Seidel iterations for

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix}$$

Construct M for both methods and find the eigenvalues of $B = I - M^{-1}A$.

- Jacobi

$$M = M^{-1} = I \rightarrow B = I - M^{-1}A = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix}$$

$\lambda = 0$ is the only eigenvalue of B.
Since $|\lambda| < 1$, Jacobi iteration converges.

- Gauss-Seidel

$$a_{11}(x_1)_{k+1} = -a_{12}(x_2)_k - a_{13}(x_3)_k + b_1$$

$$a_{22}(x_2)_{k+1} = -a_{21}(x_1)_{k+1} - a_{23}(x_3)_k + b_2$$

$$a_{33}(x_3)_{k+1} = -a_{31}(x_1)_{k+1} - a_{32}(x_2)_{k+1} + b_3$$

$$a_{11}(x_1)_{k+1} = -a_{12}(x_2)_k - a_{13}(x_3)_k + b_1$$

$$a_{22}(x_2)_{k+1} + a_{21}(x_1)_{k+1} = -a_{23}(x_3)_k + b_2$$

$$a_{33}(x_3)_{k+1} + a_{31}(x_1)_{k+1} + a_{32}(x_2)_{k+1} = b_3$$

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} x_{k+1} = \begin{bmatrix} 0 & -a_{12} & -a_{13} \\ 0 & 0 & -a_{23} \\ 0 & 0 & 0 \end{bmatrix} x_k + b$$

$$\left(M x_{k+1} = (M-A) x_k + b \right)$$

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix} \qquad M^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}$$

$$B = I - M^{-1}A = \begin{bmatrix} 0 & -2 & 2 \\ 0 & 2 & -3 \\ 0 & 0 & 2 \end{bmatrix}$$

B has eigenvalues $\lambda = 0, 2$

Since $|\lambda_2| = |2| > 1$, Gauss-Seidel diverges.

5.3.6 If $\mu = \cos \pi h$, show from (16) that
$$\lambda = \frac{1-\sin \pi h}{1+\sin \pi h} .$$
*

What is the limit of ω_{opt} as $h \to 0$? *

(16) from pg 409:

$$\omega_{opt} = \frac{2\left(1-\sqrt{1-\mu_{max}^2}\right)}{\mu_{max}^2} \qquad \lambda_{max} = \omega_{opt} - 1$$

Dropping the subscripts,

$$\omega = \frac{2\left(1-\sqrt{1-\cos^2 \pi h}\right)}{\cos^2 \pi h}$$

$$= \frac{2\left(1-|\sin \pi h|\right)}{1-\sin^2 \pi h}$$

$$= \frac{2\left(1-\sin \pi h\right)}{1-\sin^2 \pi h} \qquad h>0 \text{ small}$$

$$= \frac{2}{1+\sin \pi h} \quad \boxed{\longrightarrow 2 \text{ as } h \longrightarrow 0}$$
*

$$\lambda = \omega - 1 = \frac{2-1-\sin \pi h}{1+\sin \pi h} = \boxed{\frac{1-\sin \pi h}{1+\sin \pi h}}$$
*

<u>5.3.7</u> If λ is an eigenvalue of C with eigen-vector x, show that both $\pm\lambda$ are eigenvalues of

$$B = \begin{bmatrix} 0 & C \\ C & 0 \end{bmatrix}$$

What are the corresponding eigenvectors?

- λ is an eigenvalue of B with eigenvector $\begin{bmatrix} x \\ x \end{bmatrix}$:

$$\begin{bmatrix} 0 & C \\ C & 0 \end{bmatrix}\begin{bmatrix} x \\ x \end{bmatrix} = \begin{bmatrix} Cx \\ Cx \end{bmatrix} = \begin{bmatrix} \lambda x \\ \lambda x \end{bmatrix} = \lambda \begin{bmatrix} x \\ x \end{bmatrix}$$

- $-\lambda$ is an eigenvalue of B with eigenvector $\begin{bmatrix} x \\ -x \end{bmatrix}$:

$$\begin{bmatrix} 0 & C \\ C & 0 \end{bmatrix}\begin{bmatrix} x \\ -x \end{bmatrix} = \begin{bmatrix} -Cx \\ Cx \end{bmatrix} = \begin{bmatrix} -\lambda x \\ \lambda x \end{bmatrix} = -\lambda \begin{bmatrix} x \\ -x \end{bmatrix}$$

Gerschgorin's Circle Theorem

Every eigenvalue of B lies in at least one of the circles $C_1, \ldots, C_n$, where C_i has its center at the diagonal entry b_{ii} and its radius $r_i = \sum_{j \neq i} |b_{ij}|$ equal to the absolute sum along the rest of the row.

Proof: $Bx = \lambda x \Rightarrow \lambda x_i = \sum_j b_{ij} x_j$ for each i

$$\lambda x_i = b_{ii} x_i + \sum_{j \neq i} b_{ij} x_j$$

$$(\lambda - b_{ii}) x_i = \sum_{j \neq i} b_{ij} x_j$$

$$(\lambda - b_{ii}) x_i = \sum_{j \neq i} b_{ij} x_j \Rightarrow |\lambda - b_{ii}||x_i| \leq \sum_{j \neq i} |b_{ij}||x_j|$$

In particular for i s.t. $|x_i| \geq |x_j| \ \forall j$,

$$|\lambda - b_{ii}| \leq \sum_{j \neq i} \frac{|b_{ij}||x_j|}{|x_i|} \leq \sum_{j \neq i} |b_{ij}| = r_i$$

That is, λ lies within the circle C_i since $|\lambda - b_{ii}| \leq r_i$. Therefore λ lies in at least one of $C_1, \ldots, C_n$.

5.3.9 If a matrix B has absolute row sums less than 1,

$$|b_{i1}| + \cdots + |b_{ii}| + \cdots + |b_{in}| < 1 \text{ for each } i,$$

show from Gerschgorin's Theorem that all eigenvalues satisfy $|\lambda| < 1$.

Spse λ is an eigenvalue of B.

$|\lambda - b_{ii}| \leq r_i$ for some $i \in \{1, 2, \ldots, n\}$

$$\Rightarrow |\lambda| - |b_{ii}| \leq |\lambda - b_{ii}| \leq r_i = \sum_{j \neq i} |b_{ij}|$$

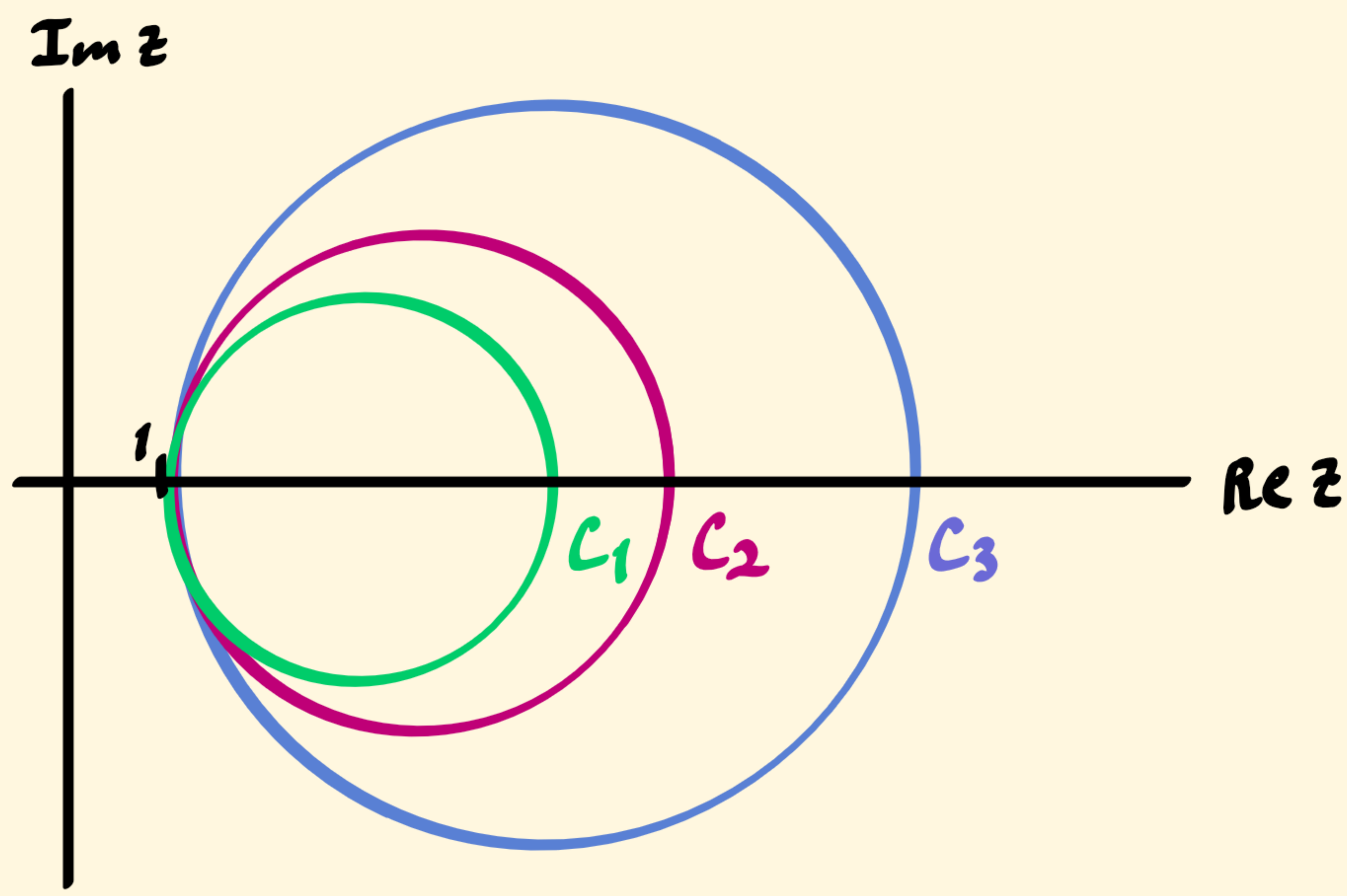
$$|\lambda| = |b_{ii}| + \sum_{j \neq i} |b_{ij}| = |b_{i1}| + \ldots + |b_{in}| < 1$$

Since λ was arbitrary, conclude $|\lambda| < 1$ for all eigenvalues λ of B.

<u>5.3.10</u> The given matrix A is diagonally dominant – each diagonal entry exceeds the sum along the rest of its row. Sketch the three Gerschgorin circles for this matrix.

$$A = \begin{bmatrix} 4 & 2 & 1 \\ 1 & 5 & 3 \\ 2 & 4 & 7 \end{bmatrix}$$

$$\begin{aligned} r_1 &= 2+1 = 3 & & C_1 = \{|z-4| < 3\} \\ r_2 &= 1+3 = 4 & \longrightarrow \quad & C_2 = \{|z-5| < 4\} \\ r_3 &= 2+4 = 6 & & C_3 = \{|z-7| < 6\} \end{aligned}$$



Note that $\lambda = 0$ is not contained in any of the circles. This holds in the general $n \times n$ case too which means diagonally dominant matrices are always invertible.

5.3.11 What is the Jacobi Matrix $B = I - D^{-1}A$ for the matrix A of 5.3.10? Show that all absolute row sums of B are less than 1 and the Jacobi iteration converges.

$$D = \begin{bmatrix} 4 & & \\ & 5 & \\ & & 7 \end{bmatrix} \longrightarrow D^{-1} = \begin{bmatrix} 1/4 & & \\ & 1/5 & \\ & & 1/7 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} - \begin{bmatrix} 1/4 & & \\ & 1/5 & \\ & & 1/7 \end{bmatrix} \begin{bmatrix} 4 & 2 & 1 \\ 1 & 5 & 3 \\ 2 & 4 & 7 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & -\frac{1}{2} & -\frac{1}{4} \\ -\frac{1}{5} & 0 & -\frac{3}{5} \\ -\frac{2}{7} & -\frac{4}{7} & 0 \end{bmatrix}$$

$$|-1/2| + |-1/4| = 3/4 < 1$$
$$|-1/5| + |-3/5| = 4/5 < 1$$
$$|-2/7| + |-4/7| = 6/7 < 1$$

By Exercise 5.3.9, each eigenvalue λ_i of B satisfies $|\lambda_i| < 1$. Then by Thm 5H, the Jacobi iteration converges.

5.3.12 Apply five steps of the unscaled power method to A starting from u_0. Do you recognize the sequence of numbers in the u_k? What is the convergence factor r and to what eigenvector do the u_k converge?

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

First five iterations:

$$u_1 = Au_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad u_2 = Au_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$u_3 = Au_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \qquad u_4 = Au_3 = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$u_5 = Au_4 = \begin{bmatrix} 8 \\ 5 \end{bmatrix}$$

Fibonacci sequence:
1, 1, 2, 3, 5, 8, 13, ...

$$0 = (1-\lambda)(-\lambda) - 1 = \lambda^2 - \lambda - 1 \rightarrow \lambda = \frac{1 \pm \sqrt{5}}{2}$$

By Thm 5I, the convergence factor is

$$r = \left|\frac{1-\sqrt{5}}{2}\right| \left|\frac{1+\sqrt{5}}{2}\right|^{-1} = \frac{\sqrt{5}-1}{1+\sqrt{5}} \approx 0.38197$$

$$0 = (1-\lambda)(-\lambda) - 1 = \lambda^2 - \lambda - 1 \rightarrow \lambda = \frac{1 \pm \sqrt{5}}{2}$$

By Thm 5I, the convergence factor is

$$r = \left|\frac{1-\sqrt{5}}{2}\right|\left|\frac{1+\sqrt{5}}{2}\right|^{-1} = \frac{\sqrt{5}-1}{1+\sqrt{5}} \approx 0.38197$$

Also by Thm 5I

$$\frac{u_k}{c_n \lambda_n^k} = \frac{c_1}{c_n}\left(\frac{\lambda_1}{\lambda_n}\right)^k x_1 + \cdots + \frac{c_{n-1}}{c_n}\left(\frac{\lambda_{n-1}}{\lambda_n}\right)^k x_{n-1} + x_n$$

$$\frac{u_k}{c_n \lambda_n^k} \longrightarrow x_n \quad \text{as} \quad k \longrightarrow \infty$$

where $u_0 = c_1 x_1 + \ldots + c_n x_n$ ← But the idea is that you don't know this decomp.?

?

The u_k diverge! However, normalizing each u_k at each iteration (scaling) gives:

```
A = [1 1; 1 0]
u_k = [1 0]';

for k = 1:20
    u_k = A*u_k;
    u_k = u_k / norm(u_k);
    if k >= 17
        u_k
    end
end

[v d] = eig(A)
```

```
u_k =

    0.85065
    0.52573

u_k =

    0.85065
    0.52573

u_k =

    0.85065
    0.52573
```

```
v =

    0.52573   -0.85065
   -0.85065   -0.52573

d =

   -0.61803          0
          0      1.618
```