Applied Complex Analysis (2021)

1 Lecture 8: Computing matrix functions via Cauchy's integral formula and the trapezium rule

In this lecture we cover the following:

1. Equivalence of the Cauchy formula to diagonalisation/Jordan canonical form.
2. Gershgorin circle theorem
3. Computing matrix functions via the trapezium rule

1.1 Equivalence of the Cauchy formula to diagonalisation/Jordan canonical form

Definition (Cauchy's integral matrix function) Suppose γ is a simple, closed contour that surrounds the spectrum of A and f is analytic in the interior. Then define

$$f(A) := \frac{1}{2\pi \mathrm{i}} \oint_{\gamma} f(\zeta)(\zeta I - A)^{-1} \mathrm{d}\zeta$$

We first show for diagonalisable f this is equivalent to the definition by diagonalisation: if $A = V\Lambda V^{-1}$ we have

$$
\begin{aligned}
\frac{1}{2\pi\mathrm{i}} \oint_\gamma f(\zeta)(\zeta I - A)^{-1}\mathrm{d}\zeta &= \frac{1}{2\pi\mathrm{i}} \oint_\gamma f(\zeta) V(\zeta I - \Lambda)^{-1} V^{-1}\mathrm{d}\zeta \\
&= \frac{1}{2\pi\mathrm{i}} V \oint_\gamma f(\zeta)(\zeta I - \Lambda)^{-1}\mathrm{d}\zeta V^{-1} \\
&= V \begin{pmatrix} \frac{1}{2\pi\mathrm{i}} \oint_\gamma f(\zeta)(\zeta - \lambda_1)^{-1}\mathrm{d}\zeta & & \\ & \ddots & \\ & & \frac{1}{2\pi\mathrm{i}} \oint_\gamma f(\zeta)(\zeta - \lambda_d)^{-1}\mathrm{d}\zeta \end{pmatrix} \mathrm{d}\zeta V^{-1} \\
&= V \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_d) \end{pmatrix} V^{-1} = f(A).
\end{aligned}
$$

For Jordan canonical form we need only show it's valid on Jordan blocks. Note that provided $\alpha \neq 0$ we have

$$\begin{pmatrix} \alpha & -1 & & \\ & \ddots & \ddots & \\ & & \alpha & -1 \\ & & & \alpha \end{pmatrix}^{-1} = \begin{pmatrix} \alpha^{-1} & \alpha^{-2} & \cdots & \alpha^{-d} \\ & \ddots & \ddots & \vdots \\ & & \alpha^{-1} & \alpha^{-2} \\ & & & \alpha^{-1} \end{pmatrix}$$

which is verifiable by inspection. Therefore for a Jordan block

$$A = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}$$

we have

$$\begin{aligned} (\zeta I - A)^{-1} &= \begin{pmatrix} \zeta - \lambda & -1 & & \\ & \ddots & \ddots & \\ & & \zeta - \lambda & -1 \\ & & & \zeta - \lambda \end{pmatrix}^{-1} \\ &= \begin{pmatrix} (\zeta - \lambda)^{-1} & (\zeta - \lambda)^{-2} & \cdots & (\zeta - \lambda)^{-d} \\ & \ddots & \ddots & \vdots \\ & & (\zeta - \lambda)^{-1} & (\zeta - \lambda)^{-2} \\ & & & (\zeta - \lambda)^{-1} \end{pmatrix} \end{aligned}$$

$$
\begin{aligned}
\frac{1}{2\pi \mathrm{i}}\oint_\gamma f(\zeta)(\zeta I - A)^{-1}\mathrm{d}\zeta &= \frac{1}{2\pi \mathrm{i}}\oint_\gamma f(\zeta)\begin{pmatrix} \zeta-\lambda & -1 & & \\ & \ddots & \ddots & \\ & & \zeta-\lambda & -1 \\ & & & \zeta-\lambda \end{pmatrix}^{-1}\mathrm{d}\zeta \\
&= \frac{1}{2\pi \mathrm{i}}\oint_\gamma f(\zeta)\begin{pmatrix} (\zeta-\lambda)^{-1} & (\zeta-\lambda)^{-2} & \cdots & (\zeta-\lambda)^{-d} \\ & \ddots & \ddots & \vdots \\ & & (\zeta-\lambda)^{-1} & (\zeta-\lambda)^{-2} \\ & & & (\zeta-\lambda)^{-1} \end{pmatrix}\mathrm{d}\zeta \\
&= \begin{pmatrix} f(\lambda) & f'(\lambda) & \cdots & f^{(d-1)}(\lambda)/(d-1)! \\ & \ddots & \ddots & \vdots \\ & & f(\lambda) & f'(\lambda) \\ & & & f(\lambda) \end{pmatrix} = f(A).
\end{aligned}
$$

1.2 Gershgorin circle theorem

If we only know A, how do we know how big to make the contour? Gershgorin's circle theorem gives the answer:

Theorem (Gershgorin) Let $A \in \mathbb{C}^{d\times d}$ and define

$$R_k = \sum_{\substack{j=1 \\ j\neq k}}^{d} |a_{kj}|$$

Then

$$\sigma(A) \subset \bigcup_{k=1}^{d} \bar{B}(a_{kk}, R_k)$$

where $\bar{B}(z_0, r)$ is the closed disk of radius r centred at z_0 and $\sigma(A)$ is the set of eigenvalues.

Proof

We can assume any eigenvalue has at least one nonzero eigenvector, whose maximum entry is 1 in the k-th entry, for some $1 \le k \le d$. (Otherwise, rescale.) That is, there exists

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_{k-1} \\ 1 \\ v_{k+1} \\ \vdots \\ v_d \end{pmatrix}$$

so that

$$A\mathbf{v} = \lambda \mathbf{v}$$

The result follows from:

$$\lambda = \mathbf{e}_k^\top (\lambda \mathbf{v}) = \mathbf{e}_k^\top A \mathbf{v} = a_{kk} + \sum_{j \ne k} a_{kj} v_j$$

so that

$$|\lambda - a_{kk}| \le \sum_{j \ne k} |a_{kj}| = R_k.$$

Demonstration Here we apply this to a particular matrix:

```
using LinearAlgebra, Plots, ComplexPhasePortrait, ApproxFun
A = [1 2 3; 1 5 2; -4 1 6]
```

```
3×@*(3 Array(*@{Int64,2}:
  1  2  3
  1  5  2
 -4  1  6
```

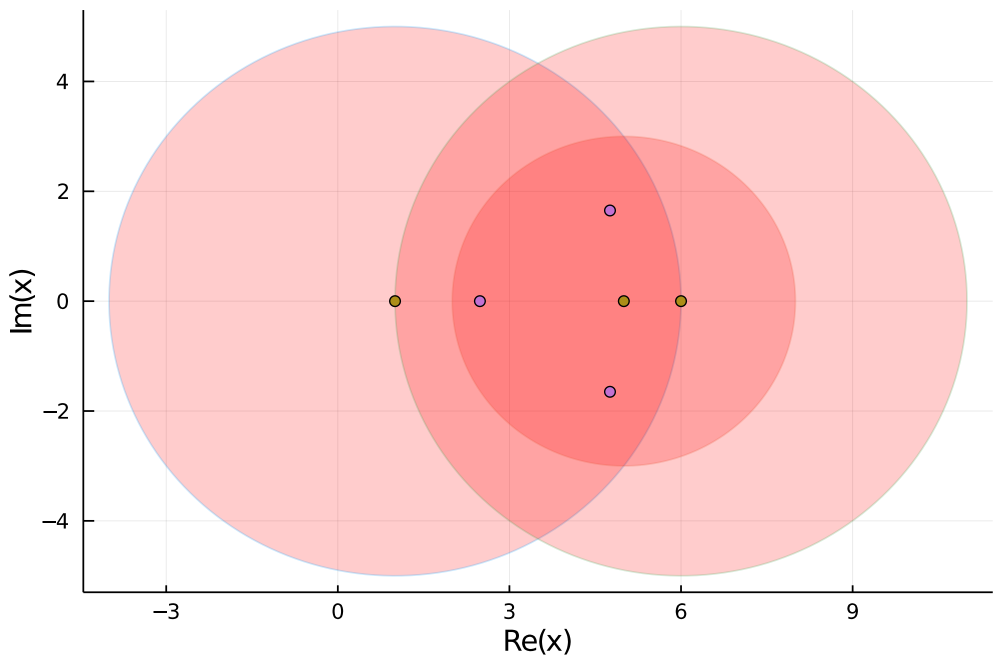
The following calculates the row sums:

```
R = sum(abs.(A - Diagonal(diag(A))),dims=2)
```

```
3×@*(1 Array(*@{Int64,2}:
 5
 3
 5
```

Gershgorin's theorem tells us that the spectrum lies in the union of the circles surrounding the diagonals:

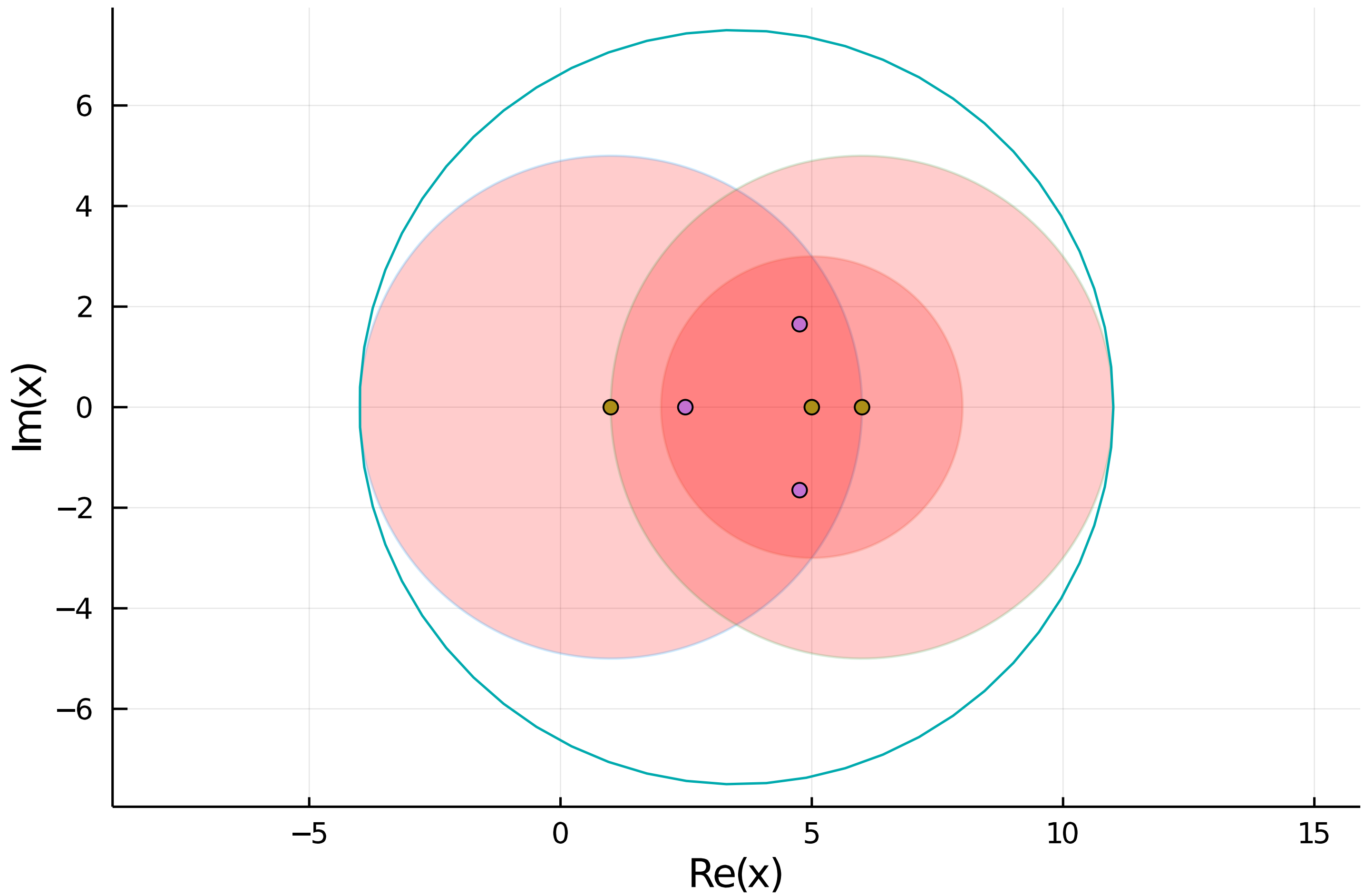
```
drawcircle!(z0, R) = plot!(θ-> real(z0) + R[1]*cos(θ), θ-> imag(z0)
+ R[1]*sin(θ), 0, 2π, fill=(0,:red), α = 0.2, legend=false)
λ = eigvals(A)
p = plot()
for k = 1:size(A,1)
    drawcircle!(A[k,k], R[k])
end
scatter!(complex.(λ); label="eigenvalues")
scatter!(complex.(diag(A)); label="diagonals")
p
```



We can therefore use this to choose a contour big enough to surround all the circles. Here's a fairly simplistic construction for our case where everything is real:

```
z_0 = (maximum(diag(A) .+ R) + minimum(diag(A) .- R)) /2 # average edges of circle
r = max(abs.(diag(A) .- R .- z_0)..., abs.(diag(A) .+ R .- z_0)...)

plot!(Circle(z_0, r);ratio=1.0)
```



1.3 Computing matrix functions via the trapezium rule

We can compute matrix functions via discretising Cauchy's integral formula with the Trapezium rule. We integrate over a simple, closed contour γ (often a circle or ellipse) that encloses the spectrum of the matrix, provided the function is analytic on and inside the contour. We can obtain such a contour from Gershgorin's circle theorem.

On the curve $\gamma : [0, 2\pi) \to \mathbb{C}$, we apply the Trapezium rule:

$$
\begin{aligned}
f(A) &= \frac{1}{2\pi \mathrm{i}} \oint_\gamma f(\zeta)(\zeta I - A)^{-1} \mathrm{d}\zeta \\
&= \frac{1}{2\pi \mathrm{i}} \int_0^{2\pi} f(\gamma(\theta))(\gamma(\theta) I - A)^{-1} \gamma'(\theta) \mathrm{d}\theta \\
&\approx \frac{1}{\mathrm{i}N} \sum_{j=0}^{N-1} f(\gamma(\theta_j))\gamma'(\theta_j)(\gamma(\theta_j) I - A)^{-1}.
\end{aligned}
$$

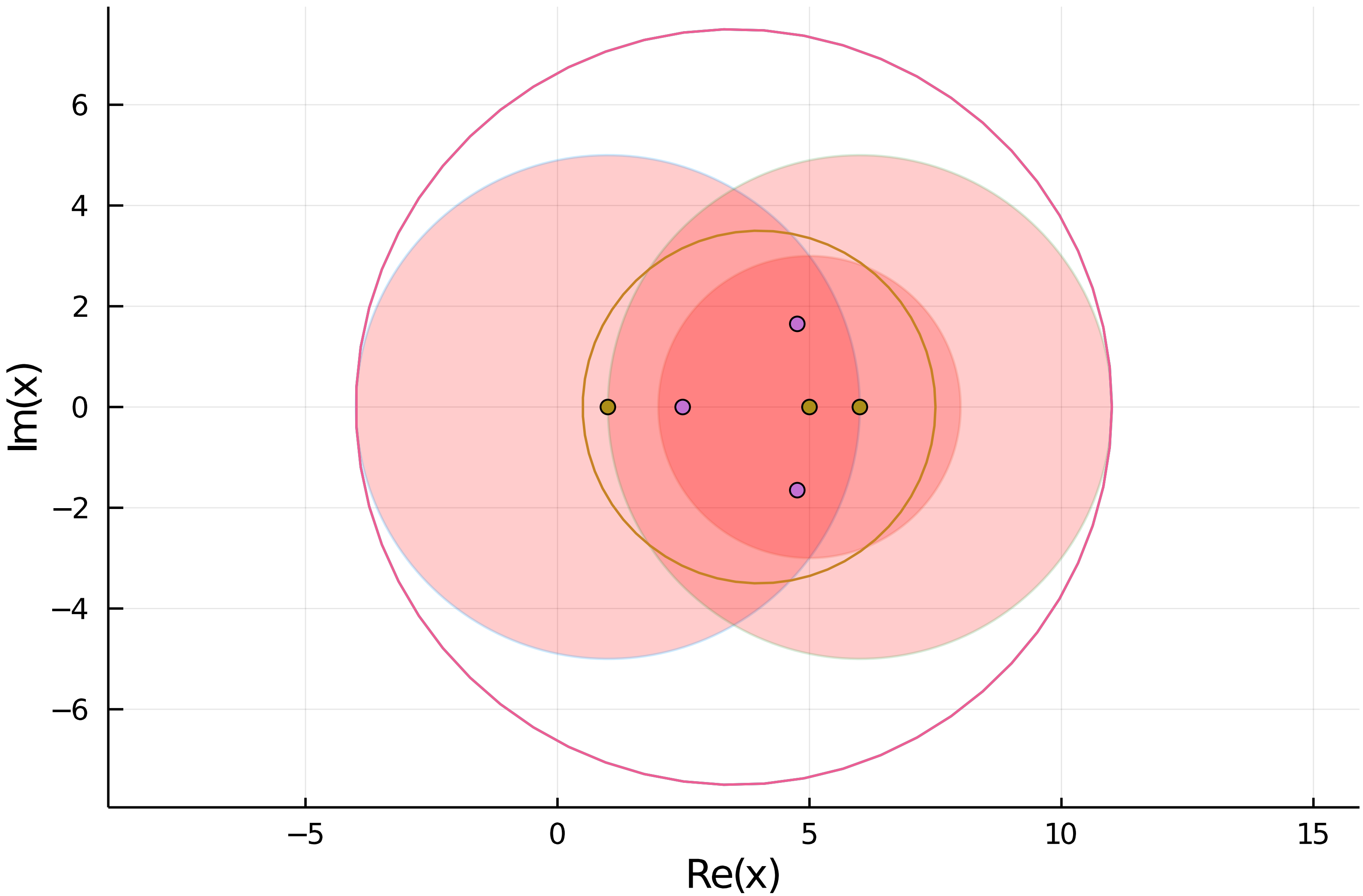
Thus matrix functions are reduced to a sum of inverses. This is useful if applying an inverse is fast, for example, we have

$$
f(A)\mathbf{v} \approx \frac{1}{\mathrm{i}N} \sum_{j=0}^{N-1} f(\gamma(\theta_j))\gamma'(\theta_j)(\gamma(\theta_j) I - A)^{-1}\mathbf{v}
$$

and if A is sparse then each inverse is fast.

Demonstration Let's compute $f(A) = A^{-1}$ via discretising Cauchy's integral formula. Consider the matrix from before, the contour we chose based on Gershgorin's theorem and a smaller contour that also encloses the spectrum of A:

```
plot!(Circle(z_0, r))
plot!(Circle(4, 3.5))
```



The Cauchy integral formula for $f(A) = A^{-1}$ is not valid if the integration contour γ is the Gershgorin contour because f is not analytic inside it; if γ is the smaller contour, then the Cauchy formula is valid:

```
periodic_rule(N) = 2π/N*(0:(N-1)), 2π/N*ones(N)
γ = θ -> z_0 + r*exp(im*θ)
γp = θ -> im*r*exp(im*θ)
N = 256
θ,w = periodic_rule(N)
norm(sum(w .* γp.(θ).* 1 ./(γ.(θ)) .* [inv(γ(θ)*I-A) for θ in
θ])/(2π*im) - inv(A))
```

```
0.7165748152568706
```

```
r = 3.5; z_0 = 4
γ = θ -> z_0 + r*exp(im*θ)
γp = θ -> im*r*exp(im*θ)
norm(sum(w .* γp.(θ).* 1 ./(γ.(θ)) .* [inv(γ(θ)*I-A) for θ in
θ])/(2π*im) - inv(A))
```

```
1.041871828445214e-15
```