Peeter Joot
peeter.joot@gmail.com

ECE1254H Modeling of Multiphysics Systems. Lecture 9: Conjugate gradient methods. Taught by Prof. Piero Triverio

1.1 Disclaimer

Peeter's lecture notes from class. These may be incoherent and rough.

1.2 Conjugate gradient convergence

For $k \ll n$, convergence orders are

We define a matrix norm, similar to $\left\lVert \mathbf{x} \right\rVert = \sqrt{\mathbf{x}^\text{T} \mathbf{x}}$, with

$$\left\lVert \mathbf{x} \right\rVert_M \equiv \sqrt{\mathbf{x}^\text{T} M \mathbf{x}}. \tag{1.1}$$

Note that our use of this is for CG which only applies to positive definite matrices (or it will not converge), so this norm is real valued.

... lots on slides...

$$K(M) = \frac{\sigma_{\text{max}}}{\sigma_{\text{min}}} \tag{1.2}$$

...

We need some fast ways to estimate the conditioning number.

1.3 Gershgorin circle theorem

Table 1.1: Convergence

	Full
Direct	$O(n^3)$
C.G.	$O(kn^2)$

Theorem 1.1: Gershgorin circle theorem

Given M, for any eigenvalue of M there is an $i \in [1, n]$ such that

$$|\lambda - M_{ii}| \le \sum_{j \ne i} |M_{ij}|$$

Consider this in the complex plane for row i

$$\begin{bmatrix} M_{i1} & M_{i2} & \cdots M_{ii} & \cdots M_{in} \end{bmatrix} \tag{1.3}$$

This inequality covers a circular region in the complex plane as illustrated in fig. 1.1 for a two eigenvalue system.

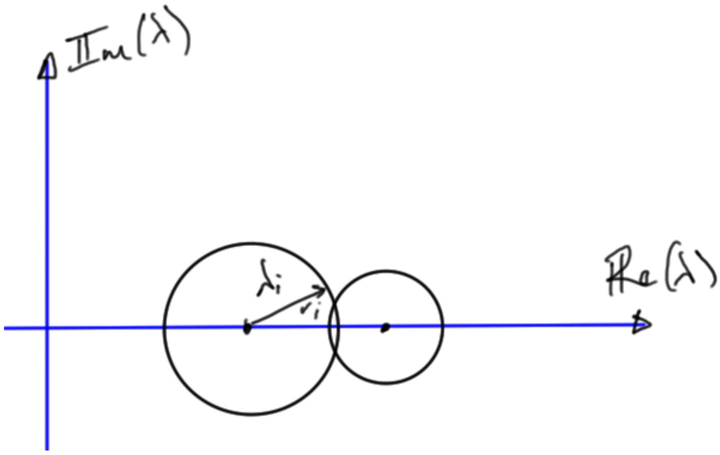

Figure 1.1: Gershgorin circles

These are called Gershgorin circles .

Example 1.1: Leaky bar

For the leaky bar of fig. 1.2, our matrix is

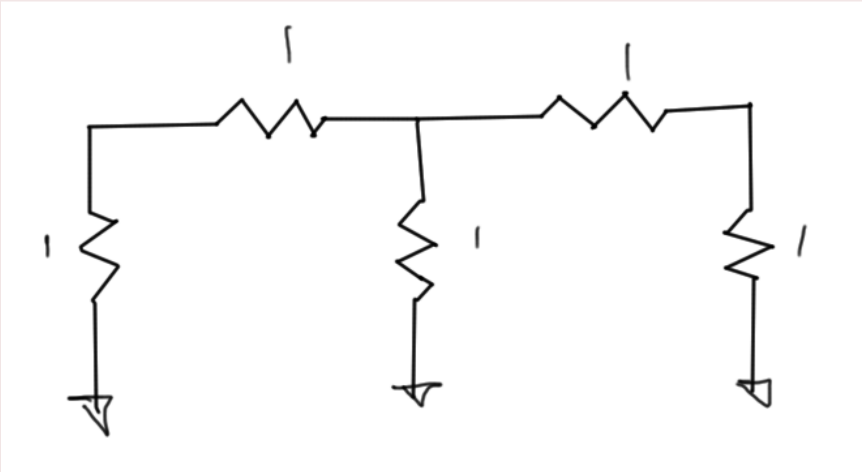

Figure 1.2: Leaky bar

$$M = \begin{bmatrix} 2 & -1 & & & \\ -1 & 3 & -1 & & \\ & -1 & 3 & -1 & \\ & & & \ddots & \\ & & & & -1 \\ & & & -1 & 2 \end{bmatrix} \tag{1.4}$$

Our Gershgorin circles are fig. 1.3.

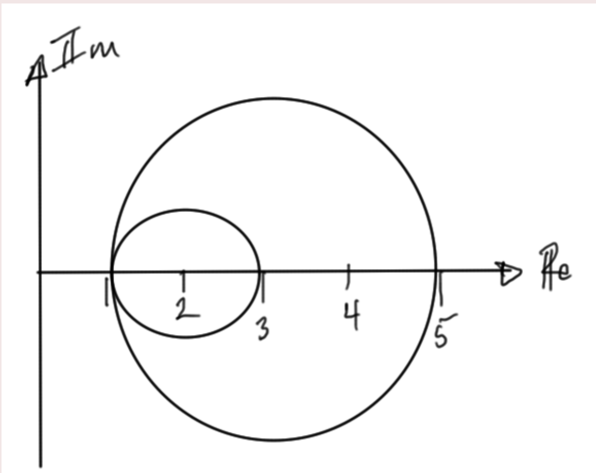

Figure 1.3: Gershgorin circles for leaky bar

We must then have

$$1 \leq \lambda(M) \leq 5, \tag{1.5}$$

so that

$$K(M) = \frac{\lambda_{\mathrm{max}}}{\lambda_{\mathrm{max}}} \leq 5. \tag{1.6}$$

On slides: example with smaller leakage to ground. On slides: example with no leakage to ground.

These had, progressively larger and larger (possibly indefinite for the latter) conditioning number estimates.

The latter had the form of

$$M = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \tag{1.7}$$

The exact eigenvalues for this system happens to be

$$\lambda \in \{3.10690.2833, 1.3049 \pm 0.7545 i\} \tag{1.8}$$

so the exact conditioning number is $3.1/0.28 \approx 11$.

Let's compare this to the estimates. Those estimates are

$$\begin{aligned} |\lambda_1 - 1| &\leq 1 \\ |\lambda_2 - 2| &\leq 2 \\ |\lambda_3 - 2| &\leq 2 \\ |\lambda_4 - 1| &\leq 1 \end{aligned} \tag{1.9}$$

These are two circles at $z = 1$ of radius 1, and two circles at $z = 2$ of radius 2, as plotted in fig. 1.4.

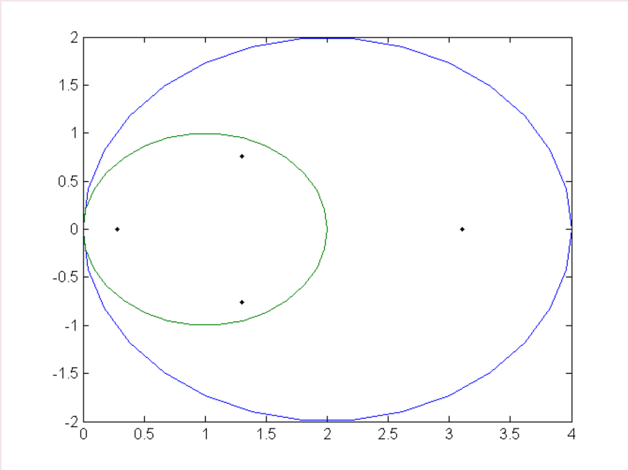

Figure 1.4: Gershgorin cirles for 4 eigenvalue system

1.4 Preconditioning

Goal is to take

$$M\mathbf{x} = \mathbf{b} \tag{1.10}$$

and introduce an easy to invert matrix P to change the problem to

$$P^{-1}M\mathbf{x} = P^{-1}\mathbf{b}. \tag{1.11}$$

This system has the same solution, but we want to choose P to maximize the convergence speed.

1.5 Symmetric preconditioning

We want to precondition in a way that preserves the symmetric positive definite nature of the matrix so that we can continue to use congugate gradient methods.

Take P and split into square root factors

$$P = P^{1/2}P^{1/2}, \tag{1.12}$$

and apply to $M\mathbf{x} = \mathbf{b}$ as

$$P^{-1/2}M\underbrace{P^{-1/2}P^{1/2}}_{=I}\mathbf{x} = P^{-1/2}\mathbf{b}. \tag{1.13}$$

and introduce a change of variables $\mathbf{y} = P^{1/2}\mathbf{x}$, so that we solve

$$P^{-1/2}MP^{-1/2}\mathbf{y} = \mathbf{b}'. \tag{1.14}$$

Some options

- Diagonal preconditioner: $\mathbf{P} = \text{diag}\{M\}$
- Incomplete LU or Cholesky factorization. Cheap, approximate decomposition where we pick $M \simeq LU = P$. An incomplete LU factorization would be easy to invert since lower or upper triangular matrices are easy to invert. In matlab we can use the ilu() function to do an incomplete LU factorization.
- ... (many preconditers are available).

For a symmetric positive definite matrix M, we can compute a LU decomposition of the form $M = LL^{\mathrm{T}}$, called the Cholesky factorization .

As an example consider the matrix

$$M = \begin{bmatrix} 2 & & & & & \text{(small)} & & \\ & 2 & & & & & & \\ & & \ddots & & & & & \\ & & & 6 & & & & \\ & & & & 6 & & & \\ & & & & & \ddots & & \\ & & & & & & 50 & \\ & & \text{(small)} & & & & & 50 \end{bmatrix}, \tag{1.15}$$

for which we can use the diagonal preconditioner

$$P = \begin{bmatrix} 1/2 & & & & & & & \\ & 1/2 & & & & & & \\ & & \ddots & & & & & \\ & & & 1/6 & & & & \\ & & & & 1/6 & & & \\ & & & & & \ddots & & \\ & & & & & & 1/50 & \\ & & & & & & & 1/50 \end{bmatrix}. \tag{1.16}$$

Our preconditioned matrix will now have the form

$$M' = \begin{bmatrix} 1 & & & & & & & \text{(small)} \\ & 1 & & & & & & \\ & & \ddots & & & & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & & & & \ddots & & \\ & & & & & & 1 & \\ \text{(small)} & & & & & & & 1 \end{bmatrix}, \tag{1.17}$$

so that the Gershgorin circles can all be found within a small radius of unity as sketched in fig. 1.5.

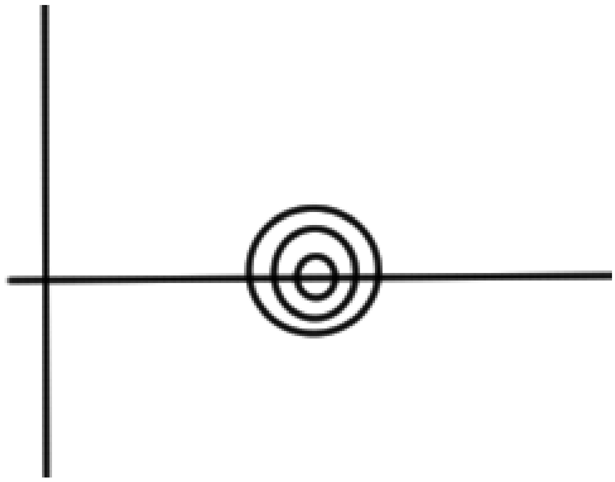

Figure 1.5: Gershgorin circles after preconditioning

1.6 Preconditioned conjugate gradient

We avoid the actually inverting the preconditioner by requiring that the LU decomposition of P be easily computed, so that we can solve

$$P\mathbf{z}^k = \mathbf{r}^k \tag{1.18}$$

by successive forward and back substitution.
More on slides...