Gerschgorin's Theorem

Let A be the matrix $A = (a_{ij})$, and let $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)'$ be an eigenvector of A corresponding to the eigenvalue λ.

For some i we have $|x_i| \geq |x_j|$ for $j \neq i$, and since $\boldsymbol{x}$ is an eigenvector, $|x_i| > 0$.

Now

$$
\begin{aligned}
A\boldsymbol{x} &= \lambda\boldsymbol{x} \\
(\lambda I - A)\boldsymbol{x} &= \mathbf{0}
\end{aligned}
$$

which represents n simultaneous equations for the x_j.

In particular, the coefficients x_j satisfy the i^{th} equation

$$
\begin{aligned}
(\lambda - a_{ii})x_i - \sum_{j\neq i} a_{ij}x_j &= 0 \\
(\lambda - a_{ii})x_i &= \sum_{j\neq i} a_{ij}x_j
\end{aligned}
$$

Therefore

$$
\begin{aligned}
|(\lambda - a_{ii})x_i| &= \left|\sum_{j\neq i} a_{ij}x_j\right| \\
|\lambda - a_{ii}|\ |x_i| &\leq \sum_{j\neq i} |a_{ij}\ |x_j| \\
&\leq \sum_{j\neq i} |a_{ij}|\,|x_i| \\
|\lambda - a_{ii}| &\leq \sum_{j\neq i} |a_{ij}|
\end{aligned}
$$

Thus the eigenvalue λ lies in one of the circles

$$
|t - a_{ii}| \leq \sum_{j\neq i} |a_{ij}|
$$

These circles are known as Gerschgorin's Circles, and this result is known as Gerschgorin's Theorem.

There are n circles corresponding to $i = 1, \ldots n$.

Suppose that $B(r)$, $0 \leq r \leq 1$ is the $n \times n$ matrix given by

$$
\begin{aligned}
b_{ii} &= a_{ii} \\
b_{ij} &= r a_{ij}\ ;\ i \neq j
\end{aligned}
$$

Then the eigenvalues of $B(r)$ lie in the circles

$$
|t - a_{ii}| \leq r \sum_{j\neq i} |a_{ij}|
$$

In particular, when $r = 0$, the eigenvalues of B are the diagonal entries a_{ij}, and there is precisely one eigenvalue in each circle.

As r increases to 1, the eigenvalues vary continuously, so that where the circles are distinct, there is one eigenvalue in the circle.

Where the circles overlap, there are the appropriate number of eigenvalues in the combined region.

For example:

If

$$A = \begin{pmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{pmatrix}$$

then the Gerschgorin circles are;

$$\begin{aligned} |t - 4| &= 1 + 0 + 0 = 1 \\ |t - 4| &= 1 + 1 + 0 = 2 \\ |t - 4| &= 0 + 1 + 1 = 2 \\ |t - 4| &= 0 + 0 + 1 = 1 \end{aligned}$$

Since these circles overlap, this merely tells us that all the eigenvalues of A satisfy $|\lambda - 4| \leq 2$.

If

$$A = \begin{pmatrix} 6 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & -2 \end{pmatrix}$$

then the Gerschgorin circles are

$$\begin{aligned} |t - 6| &= 1 \\ |t - 2| &= 2 \\ |t + 2| &= 1 \end{aligned}$$

These circles are distinct. Each circle contains one eigenvalue.

Since for real A, complex valued eigenvalues occur in conjugate pairs, both of which would lie in the same Gerschgorin circle, it follows that the eigenvalues of A are real and the dominant eigenvalue satisfies

$$5 \leq \lambda \leq 7 \ .$$

The other eigenvalues satisfy

$$\begin{aligned} 0 &\leq \lambda \leq 4 \\ -3 &\leq \lambda \leq -1 \end{aligned}$$

We can use the power method to determine the dominant eigenvalue.

Starting with $\boldsymbol{x} = (1, 0, 0)'$, we find $\lambda \sim 6.2426$ and $\boldsymbol{x} \sim (1.0000, 0.2426, 0.0294)'$

We can also use the information provided by the Gerschgorin circles to determine the least eigenvalue.

If instead of A we consider $B = 6I - A$, then the eigenvalues of B are $6 - \lambda_i$, where the eigenvalues of A are λ_i.

Therefore, corresponding to the eigenvalue of A which satisfies

$$|\lambda + 2| \leq 1$$

B has a dominant eigenvalue which satisfies

$$|\lambda - 8| \leq 1$$

The power method gives the dominant eigenvalue of B as 8.2426, so that the corresponding eigenvalue of A is -2.2426, which lies in the appropriate Gerschgorin circle. The corresponding eigenvalue is $(0.0294, -0.2426, 1.0000)'$.

We can determine the final eigenvalue (but not the eigenvector) by using the fact that

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii} (= Tr\,(A))$$

Therefore, the sum of the eigenvalues of A is $6 + 2 + (-2) = 6$.

This gives the remaining eigenvalue as

$$\lambda = 6 - (6.2426 - 2.2426) = 2\ !$$

Since the eigenvalues of A^t are the same as the eigenvalues of A, Gerschgorin's theorem applies equally well to the columns of A.

That is, the eigenvalues of A lie in the circles

$$|t - a_{ii}| \leq \sum_{j \neq i} |a_{ji}|$$

For the examples above, the matrices satisfied $A^t = A$, so that in those cases this provides no further information.

However, when $A^t \neq A$ we may be able to learn more.

For example:

$$A = \begin{pmatrix} 2 & 1 & 0 \\ \frac{1}{2} & 5 & \frac{1}{2} \\ 2 & 0 & 7 \end{pmatrix}$$

Applying Gerschgorin's Theorem to the rows, we have the circles

$$|t - 2| = 1$$
$$|t - 5| = 1$$
$$|t - 7| = 2$$

This tells us that there is one real eigenvalue with $1 \leq \lambda \leq 3$, and two possibly complex conjugate values lying in the overlapping region, with real parts between 4 and 9.

On the other hand, if we consider the rows we get the circles

$$|t-2| = 2\frac{1}{2}$$
$$|t-5| = 1$$
$$|t-7| = \frac{1}{2}$$

which tells us that there is a real eigenvalue in $[6.5, 7.5]$ and two eigenvalues in the overlapping region with real parts between $-\frac{1}{2}$ and 6.

Combining these we see that this matrix has 3 real eigenvalues which lie in the intervals $[1, 3]$, $[4, 6]$ and $[6.5, 7.5]$.

Stochastic Matrices

(Also known as *Markov* matrices)

A $n \times n$ matrix $M = (m_{ij})$ is a stochastic matrix if

(i) $$m_{ij} \geq 0 \ \forall \ i, j$$

(ii) $$\sum_{j=1}^{n} m_{ij} = 1 \ \forall i$$

Stochastic matrices arise in probability problems;
for example:

The Rentawreck company has offices in Albion, Bulimba and Chermside.

It has been found that a customer who rents from A returns the car to A 80% of the time, to B 10% of the time and to C 10% of the time.

The figures for customers who rent from B and C are
B: 30% to A, 20% to B and 50% to C;
C: 20% to A, 60% to B and 20% to C.

The stochastic matrix associated with this data is

$$M = \begin{pmatrix} 0.8 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.6 \\ 0.1 & 0.5 & 0.2 \end{pmatrix}$$

If a car has probabilities p_A, p_B and p_C, where $p_A + p_B + p_C = 1$, of starting at A, B, C respectively, we can write this as an initial vector

$$\boldsymbol{x}_0 = \begin{pmatrix} p_A \\ p_B \\ p_C \end{pmatrix}$$

At the end of the hire its probabilities of being at A, B, C are given by

$$\boldsymbol{x}_1 = M\boldsymbol{x}_0$$

and after n hires we have

$$\boldsymbol{x}_n = M^n\boldsymbol{x}_0$$

which should approximate $\lambda_1^n \boldsymbol{p}_1$ where λ_1 is the dominant eigenvalue and $\boldsymbol{p}_1$ is the corresponding eigenvector.

Applying Gerschgorin's Theorem to the columns of M, we obtain the circles

$$\begin{aligned}|t - m_{jj}| &= \sum_{i \neq j} |m_{ij}| \\ &= \sum_{i \neq j} m_{ij} \\ &= 1 - m_{ii}\end{aligned}$$

Therefore all the eigenvalues lie within the unit circle with the exception of the possible eigenvalue $\lambda = 1$.

On the other hand, if we consider M^t and the vector

$$\boldsymbol{q} = (1, 1, \ldots, 1)'$$

we see that

$$M^t \boldsymbol{q} = \begin{pmatrix} \sum_{i=1}^n m_{i1} \\ . \\ \sum_{i=1}^n m_{in} \end{pmatrix} = \boldsymbol{q}$$

so that $\lambda = 1$ is an eigenvalue of M^t and hence of M.

Therefore $M^n \boldsymbol{x}_0$ converges to an eigenvector of M corresponding to the eigenvalue $\lambda = 1$, normalised so that $||\boldsymbol{p}_1||_1 = 1$.

Problems may arise if M has a repeated eigenvalue of 1.

This happens when the problem breaks down into two (or more) disjoint problems, so that by reordering the states if necessary the matrix M has the form

$$M = \begin{pmatrix} M_1 & O \\ O & M_2 \end{pmatrix}$$

where M_1 and M_2 are smaller stochastic matrices.

However, if all of the elements of M are positive, as in the example above, this case cannot arise.

e.g for

$$M = \begin{pmatrix} 0.8 & 0.3 & 0.2 \\ 0.1 & 0.2 & 0.6 \\ 0.1 & 0.5 & 0.2 \end{pmatrix}$$

$\boldsymbol{p}_1$ satisfies

$$\begin{aligned}-0.2p_A + 0.2p_B + 0.1p_C &= 0 \\ 0.1p_A - 0.8p_B + 0.6p_C &= 0 \\ 0.1p_A + 0.5p_B - 0.8p_C &= 0\end{aligned}$$

$$-1.3p_B + 1.4p_C = 0$$

$$\begin{aligned}p_B &= 14/13p_C \\ p_A = 8p_C - 5p_B &= 34/13p_C \\ \boldsymbol{p}_1 &= c(34, 14, 13)'\end{aligned}$$

where $c^{-1} = 34 + 14 + 13 = 61$

$$\boldsymbol{p}_1 = \left(\frac{34}{61}, \frac{14}{61}, \frac{13}{61} \right)'$$

and these numbers represent the long term probabilities that a car will be at a particular depot.