GERSHGORIN'S THEOREM

GERSHGORIN'S THEOREM

Let $A=\left(a_{ij}\right)$ is an n by n matrix. Consider the union of discs $K=\bigcup K_i$ in the complex plane where

$$K_i=\left\{\mu\in C \mid \left|\mu-a_{ii}\right|\le\sum_{k=1,k\neq i}^{n}\left|a_{ik}\right|\right\}$$

is a closed disc. Then K contains all the eigenvalues of A.

This theorem is clearly true for diagonal matrix. One can think of it as a way to locate the eigenvalues when the matrix becomes not so diagonal.

Example

Consider the matrix $A = \begin{pmatrix} 1 & 0.1 & -0.1 \\ 0 & 2 & 0.4 \\ -0.2 & 0 & 3 \end{pmatrix}$, then we have

$$K_1 = \{\mu \in C \mid |\mu - 1| \le 0.2\},\ K_2 = \{\mu \in C \mid |\mu - 2| \le 0.4\},\ K_3 = \{\mu \in C \mid |\mu - 3| \le 0.2\}$$

We can do better than this: we can consider now $A^T = \begin{pmatrix} 1 & 0 & -0.2 \\ 0.1 & 2 & 0 \\ -0.1 & 0.4 & 3 \end{pmatrix}$, then we have

$$L_1 = \{\mu \in C \mid |\mu - 1| \le 0.2\},\ L_2 = \{\mu \in C \mid |\mu - 2| \le 0.1\},\ L_3 = \{\mu \in C \mid |\mu - 3| \le 0.5\}$$

If we know each disc contain exactly one eigenvalue of A, then, it is an improve as we now know each disc K_1, L_2, K_3 contain exactly one eigenvalue of A. This is true due to the following Corollary:

COROLLARY

Suppose $M_1 = \bigcup_{j=1}^{k} K_{i_j}$ be the union of k discs and the remain $n-k$ discs are disjoint, then M_1 contains exactly k eigenvalues of A, and the remaining $n-k$ discs have the property that each one contains exactly one eigenvalue of A.

Proof: Write $A = A_D + R$ where A_D is the diagonal of A. Consider the family of matrices $A_t = A_D + tR$ where $0 \le t \le 1$. Then we have $A_0 = A_D$ and $A_1 = A$. The proof is finished by noting that eigenvalues are continuous functions of t.

The theorem can be quite useful if you want to quickly argue that the matrix is non-singular, or that all of its eigevalues are less than one, as so on…

Proof of the Theorem

Let λ be an eigenvalue of A. If $\lambda = a_{ii}$ for some i there is nothing to show.

Let B be an n by n matrix .Let λ be an eigenvalue of A but not an eigenvalue of B. We first show that

$$1 \le \operatorname{lub}(\lambda I - B)^{-1}(A - B) \le \operatorname{lub}\left((\lambda I - B)^{-1}\right) \cdot \operatorname{lub}(A - B)$$

Let the vector $\overset{\vee}{v}$ be an eigenvector of A corresponding to the eigenvalues λ. Then we have $(A - B)\overset{\vee}{v} = (\lambda I - B)\overset{\vee}{v}$. So $(\lambda I - B)^{-1}(A - B)\overset{\vee}{v} = \overset{\vee}{v}$, by definition we have

$$1 \le \operatorname{lub}(\lambda I - B)^{-1}(A - B)$$

Now choose $B = A_0 = \begin{pmatrix} a_{11} & & 0 \\ & 0 & \\ 0 & & a_{nn} \end{pmatrix}$ and select the max norm $\|\vec{x}\|_\infty = \max\limits_i |x_i|$. In this case, the operator norm becomes $\operatorname{lub}(A) = \max\limits_i \left(\sum\limits_k |a_{ik}| \right)$. The expression $\operatorname{lub}(\lambda I - B)^{-1}(A - B)$ becomes $\max\limits_{1 \le i \le n} \left(\dfrac{1}{|\lambda - a_{ii}|} \sum\limits_{k=1, k \ne i}^{n} |a_{ik}| \right)$. Hence we have $1 \le \max\limits_{1 \le i \le n} \left(\dfrac{1}{|\lambda - a_{ii}|} \sum\limits_{k=1, k \ne i}^{n} |a_{ik}| \right)$ and this finishes the proof.

REFERENCES

[Hull] John Hull, *Options, Futures and other derivatives*, Prentice Hall, 6th edition

[Maj] Majewski ,*Tools for Computational Finance*, Springer, 2nd edition

[StBu] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, Springer, 3rd edition