



12.4 Positive definite/semi-definite matrices and maximum eigenvalues





Introduction

- In this topic, we will
 - Define positive definite and positive semi-definite matrices
 - Make an observation about maximum eigenvalues
 - Look at the power method for finding the maximum eigenvalue and a corresponding eigenvector





Eigenvalues

- Recall: all the eigenvalues of a self-adjoint linear operator are real
 - Thus, all the eigenvalues of a symmetric real matrix are real
 - In many cases in engineering, it can be guaranteed that all the eigenvalues of a symmetric matrix are also non-negative
- Definitions:
 - A self-adjoint linear operator that has all positive eigenvalues is described as being *positive definite*
 - A self-adjoint linear operator that has all non-negative eigenvalues is described as being positive semi-definite
- Ignoring multiplicity,
we sort the eigenvalues from largest to smallest:

$$\lambda_1 > \lambda_2 > \lambda_3 > \cdots > \lambda_{\min}$$





Positive definite matrices are invertible

- If all the eigenvalues of a matrix are positive, the matrix is invertible
 - A matrix is non-invertible or singular if and only if 0 is an eigenvalue of that matrix
- Many systems in engineering are positive definite by their very nature
 - Some systems not only are positive definite, but in addition, the largest eigenvalue is much larger than all other eigenvalues
 - That is, $\lambda_1 \gg \lambda_2 > \lambda_3 > \dots$





Finding the maximum eigenvalue

- For example, if A is any real matrix, then AA^T and $A^T A$ must both be positive semi-definite:

$$(AA^T)^T = (A^T)^T A^T = AA^T$$

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

- Now, all eigenvalues of a symmetric matrix are real

- Suppose that $(A^T A)\mathbf{u} = \lambda\mathbf{u}$

- Then $\langle \mathbf{u}, A^T A\mathbf{u} \rangle = \langle \mathbf{u}, \lambda\mathbf{u} \rangle$

$$= \lambda \langle \mathbf{u}, \mathbf{u} \rangle = \lambda \|\mathbf{u}\|_2^2$$

- However, $\langle \mathbf{u}, A^T A\mathbf{u} \rangle = \langle (A^T)^T \mathbf{u}, A\mathbf{u} \rangle$

$$= \langle A\mathbf{u}, A\mathbf{u} \rangle = \|A\mathbf{u}\|_2^2 \geq 0$$





Finding the maximum eigenvalue

- Additionally, if A is invertible, then AA^T and $A^T A$ must both be positive definite, for 0 cannot be an eigenvalue
- You can do the same with the conjugate transpose with complex matrices





Finding the maximum eigenvalue

- Question: How do we find just the maximum eigenvalue?
 - Start with a randomly generated normalized vector $\hat{\mathbf{v}}_0$
 - Because A is positive semi-definite, it is self-adjoint, so there is an orthonormal basis for $\mathbf{R}^n$
 - Thus, $\hat{\mathbf{v}}_0 = \alpha_1 \hat{\mathbf{u}}_1 + \alpha_2 \hat{\mathbf{u}}_2 + \cdots + \alpha_n \hat{\mathbf{u}}_n$

$$\begin{aligned} A\hat{\mathbf{v}}_0 &= A(\alpha_1 \hat{\mathbf{u}}_1 + \alpha_2 \hat{\mathbf{u}}_2 + \cdots + \alpha_n \hat{\mathbf{u}}_n) \\ &= \alpha_1 A\hat{\mathbf{u}}_1 + \alpha_2 A\hat{\mathbf{u}}_2 + \cdots + \alpha_n A\hat{\mathbf{u}}_n \\ &= \alpha_1 \lambda_1 \hat{\mathbf{u}}_1 + \alpha_2 \lambda_2 \hat{\mathbf{u}}_2 + \cdots + \alpha_n \lambda_n \hat{\mathbf{u}}_n \end{aligned}$$

- The coefficient that will be magnified the most is that of $\hat{\mathbf{u}}_1$





Finding the maximum eigenvalue

- Thus, if we were to continue iterating:
 - Thus, $A^k \hat{\mathbf{v}}_0 = \alpha_1 \lambda_1^k \hat{\mathbf{u}}_1 + \alpha_2 \lambda_2^k \hat{\mathbf{u}}_2 + \cdots + \alpha_n \lambda_n^k \hat{\mathbf{u}}_n$
 - Dividing by

$$\frac{A^k \hat{\mathbf{v}}_0}{\lambda_1^k} = \alpha_1 \hat{\mathbf{u}}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \hat{\mathbf{u}}_2 + \cdots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \hat{\mathbf{u}}_n$$

- Because $\lambda_1 \gg \lambda_2 > \lambda_3 > \cdots$, it follows that

$$\lim_{k \rightarrow \infty} \left(\frac{\lambda_2}{\lambda_1} \right)^k \rightarrow 0, \dots, \lim_{k \rightarrow \infty} \left(\frac{\lambda_n}{\lambda_1} \right)^k \rightarrow 0$$





Example

- Consider the following:

$$A = \begin{pmatrix} 1.01 & 0.82 & 0.82 & 1.15 & 1.27 \\ 0.82 & 1.76 & 1.02 & 0.87 & 1.37 \\ 0.82 & 1.02 & 1.3 & 1.12 & 0.94 \\ 1.15 & 0.87 & 1.12 & 1.84 & 1.53 \\ 1.27 & 1.37 & 0.94 & 1.53 & 2.31 \end{pmatrix}$$

- Its maximum eigenvalue is $\lambda_1 = 6.147187791351218$ with

$$\hat{\mathbf{u}}_1 = \begin{pmatrix} 0.372680156920819 \\ 0.425523584087080 \\ 0.371742693041126 \\ 0.480274288203895 \\ 0.557837931915653 \end{pmatrix}$$





Example

- We started with the random vector:

$$\hat{\mathbf{v}}_0 = \begin{pmatrix} -0.291597726703382 \\ 0.542998965390436 \\ -0.469118487067302 \\ -0.532149834515414 \\ 0.341858579069237 \end{pmatrix}$$





Example

- We will perform the following ten times, starting with $k = 1$:
 - Calculate $\mathbf{v}_k \leftarrow A\hat{\mathbf{v}}_{k-1}$
 - Print $\|\mathbf{v}_k\|_2$
 - Normalize $\hat{\mathbf{v}}_k \leftarrow \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|_2}$

$$\|\mathbf{v}_1\|_2 = 1.630384666995952$$

$$\|\mathbf{v}_2\|_2 = 5.395182239509930$$

$$\|\mathbf{v}_3\|_2 = 6.124595853758402$$

$$\|\mathbf{v}_4\|_2 = 6.146623014513120$$

$$\|\mathbf{v}_5\|_2 = 6.147173744687314$$

$$\|\mathbf{v}_6\|_2 = 6.147187442032004$$

$$\|\mathbf{v}_7\|_2 = 6.147187782664097$$

$$\|\mathbf{v}_8\|_2 = 6.147187791135181$$

$$\|\mathbf{v}_9\|_2 = 6.147187791345845$$

$$\|\mathbf{v}_{10}\|_2 = 6.147187791351087$$

$$\hat{\mathbf{v}}_{10} = \begin{pmatrix} -0.372680150617751 \\ -0.425523610607205 \\ -0.371742696438822 \\ -0.480274269514795 \\ -0.557837929723107 \end{pmatrix}$$

$$\hat{\mathbf{u}}_1 = \begin{pmatrix} 0.372680156920819 \\ 0.425523584087080 \\ 0.371742693041126 \\ 0.480274288203895 \\ 0.557837931915653 \end{pmatrix}$$





Diagonalization

- Thus, we started with

$$\hat{\mathbf{v}}_0 = \begin{pmatrix} -0.291597726703382 \\ 0.542998965390436 \\ -0.469118487067302 \\ -0.532149834515414 \\ 0.341858579069237 \end{pmatrix}$$

- This has an angle with the eigenspace of 76.56°
- After 10 iterations, $\hat{\mathbf{v}}_{10}$ is a close approximation to a normalized eigenvector of the eigenvalue 6.147187791351087





Comments

- There are better algorithms for finding the maximum eigenvalue and an associated eigenvector,
but this is never-the-less interesting
- If the second-largest eigenvalue is close in magnitude to the first, it seriously slows convergence
- This algorithm can be used for finding the largest eigenvalue if it is negative, but it requires additional considerations
 - It can even be used to find the largest-in-magnitude eigenvalue if it is complex, but again, now we are dealing with a rotating eigenvector





So what?

- Why do we care only about the maximum eigenvalue?
 - If $\lambda_1 \gg \lambda_2$, then generally a reasonable approximation to $A\mathbf{v}$ is

$$A\mathbf{v} \approx \lambda_1 \langle \hat{\mathbf{u}}_1, \mathbf{v} \rangle \hat{\mathbf{u}}_1$$

$$A\hat{\mathbf{e}}_1 = \begin{pmatrix} 1.01 \\ 0.82 \\ 0.82 \\ 1.15 \\ 1.27 \end{pmatrix} \quad A\hat{\mathbf{e}}_2 = \begin{pmatrix} 0.82 \\ 1.76 \\ 1.02 \\ 0.87 \\ 1.37 \end{pmatrix} \quad A\hat{\mathbf{e}}_3 = \begin{pmatrix} 0.82 \\ 1.02 \\ 1.3 \\ 1.12 \\ 0.94 \end{pmatrix} \quad A\hat{\mathbf{e}}_4 = \begin{pmatrix} 1.15 \\ 0.87 \\ 1.12 \\ 1.84 \\ 1.53 \end{pmatrix} \quad A\hat{\mathbf{e}}_5 = \begin{pmatrix} 1.27 \\ 1.37 \\ 0.94 \\ 1.53 \\ 2.31 \end{pmatrix}$$

$$\lambda_1 \langle \hat{\mathbf{u}}_1, \hat{\mathbf{e}}_1 \rangle \hat{\mathbf{u}}_1 = \begin{pmatrix} 0.854 \\ 0.975 \\ 0.852 \\ 1.100 \\ 1.278 \end{pmatrix} \quad \begin{pmatrix} 0.975 \\ 1.113 \\ 0.972 \\ 1.256 \\ 1.459 \end{pmatrix} \quad \begin{pmatrix} 0.852 \\ 0.972 \\ 0.849 \\ 1.098 \\ 1.275 \end{pmatrix} \quad \begin{pmatrix} 1.100 \\ 1.256 \\ 1.098 \\ 1.418 \\ 1.647 \end{pmatrix} \quad \begin{pmatrix} 1.278 \\ 1.459 \\ 1.275 \\ 1.647 \\ 1.913 \end{pmatrix}_{14}$$





Summary

- Following this topic, you now
 - Know what a positive definite/semi-definite matrix is
 - Know that given any matrix A , both AA^T and A^TA are positive semi-definite,
and if A square and is invertible, both are positive definite
 - Understand the power method allows us to approximate the largest eigenvalue of such a matrix
 - Are aware that the rate of convergence depends on how much larger the largest eigenvalue is in relation to the second-largest eigenvalue
 - Understand that if $\lambda_1 \gg \lambda_2$, then this largest eigenvector can be used to well approximate the behavior of the entire matrix





References

- [1] https://en.wikipedia.org/wiki/Diagonalizable_matrix#Diagonalization





Acknowledgments

None so far.





Colophon

These slides were prepared using the Cambria typeface. Mathematical equations use Times New Roman, and source code is presented using Consolas. Mathematical equations are prepared in MathType by Design Science, Inc. Examples may be formulated and checked using Maple by Maplesoft, Inc.

The photographs of flowers and a monarch butter appearing on the title slide and accenting the top of each other slide were taken at the Royal Botanical Gardens in October of 2017 by Douglas Wilhelm Harder. Please see

<https://www.rbg.ca/>

for more information.





Disclaimer

These slides are provided for the NE 112 *Linear algebra for nanotechnology engineering* course taught at the University of Waterloo. The material in it reflects the authors' best judgment in light of the information available to them at the time of preparation. Any reliance on these course slides by any party for any other purpose are the responsibility of such parties. The authors accept no responsibility for damages, if any, suffered by any party as a result of decisions made or actions based on these course slides for any other purpose than that for which it was intended.

