

Practical Linear Algebra: A GEOMETRY TOOLBOX

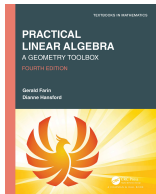
Fourth edition

Chapter 15: Eigen Things Revisited

Gerald Farin & Dianne Hansford

A K Peters/CRC Press
www.farinhanford.com/books/pla

©2021

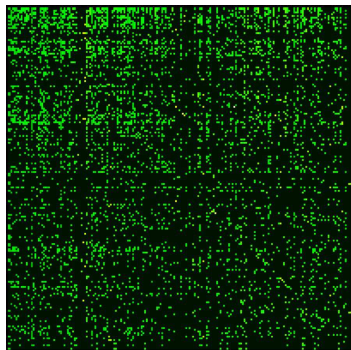


Outline

- 1 Introduction to Eigen Things Revisited
- 2 The Basics Revisited
- 3 Similarity and Diagonalization
- 4 Quadratic Forms
- 5 The Power Method
- 6 Application: Google Eigenvector
- 7 QR Algorithm
- 8 Eigenfunctions
- 9 WYSK

Introduction to Eigen Things Revisited

Eigenvalues and eigenvectors reveal action and geometry of map



Connectivity matrix for a Google matrix

Chapter 7: 2×2 matrices

Here: $n \times n$ matrices

Important in many areas:

- characterizing harmonics of musical instruments
- moderating movement of fuel in a ship
- analysis of large data sets

Google matrix: webpage ranking

The Basics Revisited

If an $n \times n$ matrix A has *fixed directions*

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

meaning that A maps \mathbf{r} to a scalar multiple of itself

$\mathbf{r} = \mathbf{0}$ trivially satisfies this equation — not interesting

Write the equation above in matrix form

$$[A - \lambda I]\mathbf{r} = \mathbf{0}$$

If $[A - \lambda I]$ maps $\mathbf{r} \neq \mathbf{0}$ to $\mathbf{0}$ then

$$p(\lambda) = \det[A - \lambda I] = 0 \quad \text{characteristic equation}$$

$p(\lambda)$ is a polynomial of degree n in λ — its zeroes are A 's **eigenvalues**

The Basics Revisited

Example:

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

$$p(\lambda) = \det[A - \lambda I] = \begin{vmatrix} 1 - \lambda & 1 & 0 & 0 \\ 0 & 3 - \lambda & 1 & 0 \\ 0 & 0 & 4 - \lambda & 1 \\ 0 & 0 & 0 & 2 - \lambda \end{vmatrix}$$

$$p(\lambda) = (1 - \lambda)(3 - \lambda)(4 - \lambda)(2 - \lambda) = 0$$

$$\text{zeroes of } p(\lambda) : \quad \lambda_1 = 4 \quad \lambda_2 = 3 \quad \lambda_3 = 2 \quad \lambda_4 = 1$$

Convention: order the eigenvalues in decreasing order

Dominant eigenvalue: largest eigenvalue in absolute value

The Basics Revisited

Not always dealing with *upper triangular matrices* like the one in the previous Example

General $n \times n$ matrix has a degree n characteristic polynomial

$$p(\lambda) = \det[A - \lambda I] = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdot \dots \cdot (\lambda_n - \lambda)$$

Finding zeroes of n^{th} degree polynomial non-trivial

Gauss elimination or LU decomposition change eigenvalues

Instead *diagonalization* can create simpler eigenvalue problems

— See Section 15.2 Similarity and Diagonalization

Iterative methods exist to find the dominant eigenvalue

— See Section 15.4 The Power Method

The Basics Revisited

Example: Elementary row operations change the eigenvalues

$$A = \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix}$$

$\det A = 2$ and eigenvalues $\lambda_1 = 2 + \sqrt{2}$ and $\lambda_2 = 2 - \sqrt{2}$

One step of forward elimination:

$$A' = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix}$$

Determinant is invariant under forward elimination: $\det A' = 2$

The eigenvalues are not: A' has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 1$

The Basics Revisited

Understand/classify eigenvalues without actually calculating them:

Characteristic equation reveals

$$p(\lambda) = \det(A - \lambda I) = (-\lambda)^n + \dots + \det A = (-\lambda)^n + \dots + \lambda_1 \lambda_2 \cdots \lambda_n$$

The *trace* of A is defined as

$$\begin{aligned}\operatorname{tr}(A) &= \lambda_1 + \lambda_2 + \dots + \lambda_n \\ &= a_{1,1} + a_{2,2} + \dots + a_{n,n}.\end{aligned}$$

Example: given a real symmetric matrix \Rightarrow real, positive eigenvalues
If trace is zero then the eigenvalues must all be zero

A and A^T have the same eigenvalues

A is invertible and has eigenvalues λ_i ; then A^{-1} has eigenvalues $1/\lambda_i$;

The Basics Revisited

Found the λ_i — now solve homogeneous linear systems

$$[A - \lambda_i I] \mathbf{r}_i = \mathbf{0}$$

to find the **eigenvectors** \mathbf{r}_i for $i = 1, n$

\mathbf{r}_i in the *null space* of $[A - \lambda_i I]$

Homogeneous systems \Rightarrow no unique solution

The solution space is called the **eigenspace** of A corresponding to λ_i

Sometimes eigenvectors normalized to eliminate ambiguity

The Basics Revisited

Example: Find the eigenvectors

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \quad \lambda_i = 4, 3, 2, 1$$

Starting with $\lambda_1 = 4$:

$$\begin{bmatrix} -3 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2 \end{bmatrix} \mathbf{r}_1 = \mathbf{0} \Rightarrow \mathbf{r}_1 = \begin{bmatrix} 1/3 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

Repeating for all eigenvalues

$$\mathbf{r}_2 = \begin{bmatrix} 1/2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{r}_3 = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ 1 \end{bmatrix} \quad \mathbf{r}_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{and check: } A\mathbf{r}_i = \lambda_i\mathbf{r}_i$$

The Basics Revisited

Multiple zeroes of the characteristic polynomial

⇒ identical homogeneous systems $[A - \lambda I]\mathbf{r} = \mathbf{0}$

Example:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \lambda_i = 2, 2, 1$$

$$\text{For } \lambda_1 = \lambda_2 = 2 \quad \begin{bmatrix} -1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{r}_1 = \mathbf{0}$$

Rank 1 matrix ⇒ 2D null space ⇒ 2D eigenspace

Two free parameters: r_3 and r_2 — Set to zero or one ...

$$\mathbf{r}_1 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{r}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

Any linear combination of these eigenvectors

The Basics Revisited

Example continued

$$\text{For } \lambda_3 = 1 \quad \begin{bmatrix} 0 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{r}_3 = \mathbf{0} \quad \Rightarrow \quad \mathbf{r}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Construct linearly independent eigenvectors in different eigenspaces
then merge them to form the complete set of eigenvectors

This complete set will be linearly independent

The Basics Revisited

Example: Rotation around the \mathbf{e}_3 -axis:

$$A = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Expect that \mathbf{e}_3 is an eigenvector:

$$A\mathbf{e}_3 = \mathbf{e}_3 \Rightarrow \text{corresponding eigenvalue} = 1$$

The Basics Revisited

Symmetric matrix A :

- real eigenvalues
- eigenvectors are orthogonal
- diagonalizable:

Diagonal matrix $\Lambda = R^{-1}AR$

Called the **eigendecomposition**

Columns of R holds eigenvectors

Λ holds eigenvalues

The Basics Revisited

Example: Eigendecomposition $\Lambda = R^{-1}SR$ of the symmetric matrix

$$S = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 3 \end{bmatrix} \quad \lambda_i = 4, 3, 2$$

Corresponding eigenvectors

$$\mathbf{r}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{r}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{r}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\Lambda = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad R = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}$$

The Basics Revisited

Projection matrices:

— eigenvalues are one or zero

0: eigenvector projected to the zero vector

⇒ determinant is zero and matrix is singular

1: eigenvector projected to itself

— If $\lambda_1 = \dots = \lambda_k = 1$ then eigenvectors populate column space

⇒ dimension is k and null space is dimension $n - k$

The Basics Revisited

Example: 3×3 projection matrix $P = \mathbf{u}\mathbf{u}^T$

$$\mathbf{u} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \quad P = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}$$

$$\lambda_1 = 1 \quad \lambda_2 = 0 \quad \lambda_3 = 0$$

$$\lambda_1 = 1 \Rightarrow \begin{bmatrix} -1/2 & 0 & 1/2 \\ 0 & -1 & 0 \\ 1/2 & 0 & -1/2 \end{bmatrix} \mathbf{r}_1 = \mathbf{0} \Rightarrow \mathbf{r}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

The Basics Revisited

Example: continued

$$\lambda_2 = \lambda_3 = 0 \quad \Rightarrow \quad \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix} \mathbf{r} = \mathbf{0}$$

Find two eigenvectors that span 2D null space:

— free parameters r_2 and r_3

$$\mathbf{r}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \mathbf{r}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

All linear combinations

$$\mathbf{r} = c\mathbf{r}_2 + d\mathbf{r}_3$$

span the eigenspace corresponding to $\lambda_2 = 0$

Use Gram-Schmidt to form an orthogonal set of vectors

Similarity and Diagonalization

Examine change of basis as a tool for transforming a matrix to a simpler one with the same eigenvalues

— Simpler matrix will be a diagonal matrix containing the eigenvalues

A : **a**-basis \rightarrow **e**-basis A^{-1} : **e**-basis \rightarrow **a**-basis

If M is a linear map in the **e**-basis then

$$M' = A^{-1}MA \quad \text{linear map in the } \mathbf{a} \text{ - basis}$$

Matrices M and M' are **similar**

— Share the same eigenvalues

— Do not share the same eigenvectors

Similarity and Diagonalization

Example: change of basis for projecting a point onto a line (Example 5.10)

Solution: $M_2 = R_\theta P R_{-\theta}$

— rotate into \mathbf{e} -basis via $R_{-\theta}$

— apply projection P

— reverse rotation via R_θ

Projection matrix P is *similar* to M_2

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \quad R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

Similarity and Diagonalization

Matrix M is **diagonalizable** if there exists an invertible matrix R such that

$$\Lambda = R^{-1}MR \quad \text{is a diagonal matrix}$$

- Eigenvalues of M are the diagonal entries in Λ
- Eigenvectors of M are the column vectors of R
Called the **eigenbasis**

Special: **orthogonally diagonalizable**

$$\Lambda = R^TMR \quad \text{orthogonal } R$$

$$\Rightarrow M \text{ is symmetric}$$

Similarity and Diagonalization

Example: distinct eigenvalues not necessary to be diagonalizable

$$\Lambda = R^T M R$$

$$M = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad R = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad \Lambda = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$

If an eigenvalue of a symmetric matrix is repeated k times
then the eigenspace spanned by the eigenvectors is k -dimensional

Gram-Schmidt method applied to achieve an orthonormal basis

\Rightarrow a symmetric matrix will always be diagonalizable

Similarity and Diagonalization

Matrix does not have to be symmetric to be diagonalizable

$n \times n$ matrix M

Diagonalizable $\Lambda = R^T M R \equiv n$ linearly independent eigenvectors

— Matrix R is not unique; recall the free parameters

Similarity and Diagonalization

Not all matrices are diagonalizable

Example: shear

$$M = \begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix} \quad \text{where } \lambda_1 = \lambda_2 = 1$$

has just one fixed direction

Similarity and Diagonalization

If a matrix M is diagonalizable matrix then repeating the map is simple

$$M = R\Lambda R^{-1} \quad \Rightarrow \quad M^k = R\Lambda^k R^{-1}$$

$$\Lambda^k = \begin{bmatrix} (\lambda_1)^k & & & \\ & \ddots & & \\ & & \ddots & \\ & & & (\lambda_n)^k \end{bmatrix}$$

This topic will be the focus of the power method

Quadratic Forms

Quadratic forms in \mathbb{R}^n

Given: \mathbf{v} in \mathbb{R}^n and $n \times n$ symmetric matrix A

$$f(\mathbf{v}) = \mathbf{v}^T A \mathbf{v} = a_{1,1}v_1^2 + 2a_{1,2}v_1v_2 + \dots + a_{n,n}v_n^2$$

Quadratic Forms

Examples:

$$A_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A_2 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A_3 = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Quadratic form for each matrix:

$$f_1(\mathbf{v}) = 2v_1^2 + 0.5v_2^2 + v_3^2 \quad f_2(\mathbf{v}) = 2v_1^2 + v_3^2 \quad f_3(\mathbf{v}) = -2v_1^2 + 0.5v_2^2 + v_3^2$$

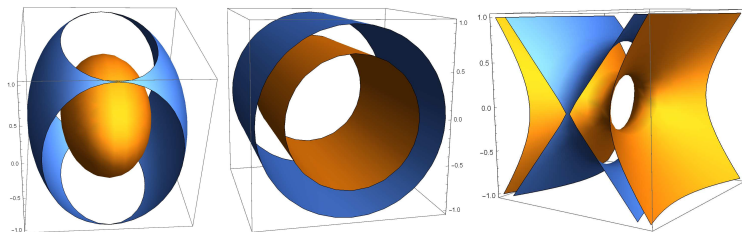
$$|A_1| = 1 \quad \lambda_1 = 2 \quad \lambda_2 = 1 \quad \lambda_3 = 0.5$$

$$|A_2| = 0 \quad \lambda_1 = 2 \quad \lambda_2 = 1 \quad \lambda_3 = 0$$

$$|A_3| = -1 \quad \lambda_1 = -2 \quad \lambda_2 = 1 \quad \lambda_3 = 0.5$$

Quadratic Forms

Contour plots



Left: $f_1 = 0.5$ and $f_1 = 1.5$

— Outer ellipsoid clipped against the bounding box

Middle: $f_2 = 0.5$ and $f_2 = 1$

Right: $f_3 = 0.1$ and $f_3 = 0.5$

Quadratic Forms

Positive definite matrix: a real matrix satisfying

$$f(\mathbf{v}) = \mathbf{v}^T \mathbf{A} \mathbf{v} > 0 \quad \text{for any } \mathbf{v} \neq \mathbf{0} \in \mathbb{R}^n$$

\Rightarrow quadratic form is positive everywhere except for $\mathbf{v} = \mathbf{0}$

Contour $f(\mathbf{v}) = 1$ is an n -dimensional ellipsoid

- Semi-minor axis corresponds to \mathbf{r}_1 with length $1/\sqrt{\lambda_1}$
- Semi-major axis corresponds to \mathbf{r}_n with length $1/\sqrt{\lambda_n}$

Example: Only A_1 is positive definite

$f_1(\mathbf{v}) = 1$ ellipsoid:

- Shortest axis length $1/\sqrt{2} = 0.5$ along the \mathbf{e}_1 -axis
- Longest axis length $1/\sqrt{0.5} = 2$ along the \mathbf{e}_2 -axis

Quadratic Forms

Rayleigh quotient $q(\mathbf{v}) = \frac{\mathbf{v}^T A \mathbf{v}}{\mathbf{v}^T \mathbf{v}}$

Maximum in the dominant eigenvector direction: $q(\mathbf{r}_1) = \lambda_1$

Minimum in direction corresponding to the smallest eigenvalue: $q(\mathbf{r}_n) = \lambda_n$

Rayleigh quotient used to approximate eigenvalues and eigenvectors
— the power method

The Power Method

A : symmetric $n \times n$ matrix with n distinct eigenvalues

Let λ be the *dominant eigenvalue* and \mathbf{r} its corresponding eigenvector

$$A^i \mathbf{r} = \lambda^i \mathbf{r}$$

Use this property to find the dominant eigenvalue and eigenvector

The Power Method

Choose an arbitrary vector $\mathbf{r}^{(1)}$ (non-zero/ $\alpha_1 \neq 0$)

There exists $\alpha_1, \alpha_2, \dots, \alpha_n$ such that

$$\mathbf{r}^{(1)} = \alpha_1 \mathbf{r}_1 + \alpha_2 \mathbf{r}_2 + \dots + \alpha_n \mathbf{r}_n$$

Apply a linear map:

$$\begin{aligned} A\mathbf{r}^{(1)} &= A(\alpha_1 \mathbf{r}_1 + \alpha_2 \mathbf{r}_2 + \dots + \alpha_n \mathbf{r}_n) \\ &= \alpha_1 \lambda_1 \mathbf{r}_1 + \alpha_2 \lambda_2 \mathbf{r}_2 + \dots + \alpha_n \lambda_n \mathbf{r}_n \end{aligned}$$

Repeat

$$A^i \mathbf{r}^{(1)} = \lambda_1^i \left(\alpha_1 \mathbf{r}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1} \right)^i \mathbf{r}_2 + \dots + \alpha_n \left(\frac{\lambda_n}{\lambda_1} \right)^i \mathbf{r}_n \right)$$

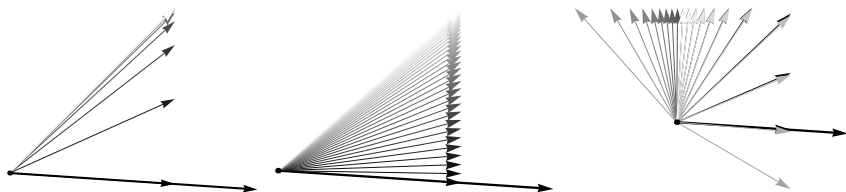
For large i $A^i \mathbf{r}^{(1)} \approx \alpha_1 \lambda_1^i \mathbf{r}_1$

The Power Method

Instead of computing A^i directly, iteratively apply A

$$\mathbf{r}^{(k+1)} = A\mathbf{r}^{(k)} \approx A(\alpha_1\lambda_1^k\mathbf{r}_1) = \lambda_1(\alpha_1\lambda_1^k\mathbf{r}_1) = \lambda_1\mathbf{r}^{(k)} \quad k = 1, 2, \dots$$

After a sufficiently large k : $\mathbf{r}^{(k)}$ will begin to line up with \mathbf{r}_1



Vector sequences examples:

- Initial guess: longest black vector
- Successive iterations: lighter gray
- Left and middle figures demonstrate convergence
- (Each iteration scaled with ∞ -norm)

The Power Method

Find λ_1 :

$$(\mathbf{r}^{(k)})^T \mathbf{r}^{(k+1)} \approx \lambda_1 (\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}$$

$$\frac{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k+1)}}{(\mathbf{r}^{(k)})^T \mathbf{r}^{(k)}} \approx \lambda_1 \quad \text{Rayleigh quotient}$$

In the algorithm to follow all components of $\mathbf{r}^{(k+1)}$ and $\mathbf{r}^{(k)}$ are (approximately) related by

$$\frac{r_j^{(k+1)}}{r_j^{(k)}} = \lambda_1 \quad \text{for } j = 1, \dots, n$$

Rather than checking each ratio, use the ∞ -norm to define λ_1 upon each iteration

The Power Method

Algorithm:

Initialization:

Estimate dominant eigenvector $\mathbf{r}^{(1)} \neq \mathbf{0}$

Find j where $|r_j^{(1)}| = \|\mathbf{r}^{(1)}\|_\infty$ and set $\mathbf{r}^{(1)} = \mathbf{r}^{(1)}/r_j^{(1)}$

Set $\lambda^{(1)} = 0$

Set tolerance ϵ

Set maximum number of iterations m

For $k = 2, \dots, m$

$\mathbf{y} = A\mathbf{r}^{(k-1)}$

$\lambda^{(k)} = y_j$

Find j where $|y_j| = \|\mathbf{y}\|_\infty$

If $y_j = 0$ Then output: "eigenvalue zero; select new $\mathbf{r}^{(1)}$ and restart"; exit

$\mathbf{r}^{(k)} = \mathbf{y}/y_j$

If $|\lambda^{(k)} - \lambda^{(k-1)}| < \epsilon$ Then output: $\lambda^{(k)}$ and $\mathbf{r}^{(k)}$; exit

If $k = m$ output: maximum iterations exceeded

The Power Method

Some remarks on this method:

- If $|\lambda|$ is either “large” or “close” to zero, could cause numerical problems — Good to *scale* the $\mathbf{r}^{(k)}$ — Done here with ∞ -norm
- Convergence seems impossible if $\mathbf{r}^{(1)}$ is perpendicular to \mathbf{r} , but numerical round-off helps and it will converge slowly
- Very slow convergence if $|\lambda_1| \approx |\lambda_2|$
- Limited to symmetric matrices with one dominant eigenvalue
May be generalized to more cases

Example application of the ∞ -norm

$$\text{guess } \mathbf{r}^{(1)} = \begin{bmatrix} 1.5 \\ -0.1 \end{bmatrix} \quad \infty\text{-norm scaled} \quad \Rightarrow \quad \mathbf{r}^{(1)} = \begin{bmatrix} 1 \\ -0.066667 \end{bmatrix}$$

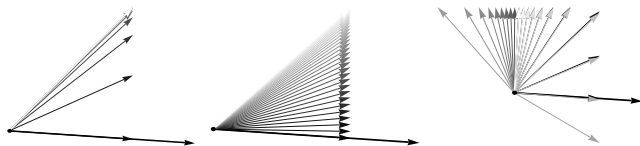
The Power Method

Example:

$$A_1 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \lambda_1 = 3 \quad \lambda_2 = 1$$

$$A_2 = \begin{bmatrix} 2 & 0.1 \\ 0.1 & 2 \end{bmatrix} \quad \lambda_1 = 2.1 \quad \lambda_2 = 1.9$$

$$A_3 = \begin{bmatrix} 2 & -0.1 \\ 0.1 & 2 \end{bmatrix} \quad \lambda_1 = 2 + 0.1i \quad \lambda_2 = 2 - 0.1i$$



Power method demonstration with A_1, A_2, A_3 (left to right)

- Initial guess: longest black vector
- Successive iterations: lighter gray
- Each iteration scaled with ∞ -norm

The Power Method

Example: continued

A_1 : symmetric and λ_1 separated from λ_2

\Rightarrow rapid convergence in 11 iterations — Estimate: $\lambda = 2.99998$

A_2 : symmetric but λ_1 close to λ_2

\Rightarrow convergence slower 41 iterations — Estimate: $\lambda = 2.09549$

A_3 : rotation matrix (not symmetric) and complex eigenvalues

\Rightarrow no convergence.

Application: Google Eigenvector

Linear algebra + search engines

Search engine techniques are highly proprietary and changing

All share the basic idea of *ranking* webpages

Concept introduced by Brin and Page in 1998 — Google

Ranking webpages is an eigenvector problem!

The web frozen at some point in time consists of N webpages

— A page pointed to very often: important

— A page with none or few other pages pointing to it: unimportant

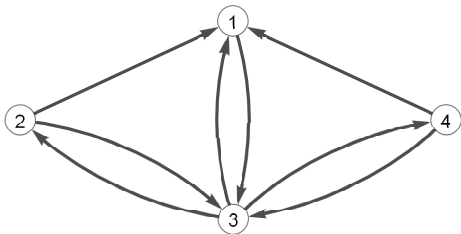
How can we rank all web pages?

Application: Google Eigenvector

Basics:

- Assume all webpages are ordered: assign a number i to each
- If page $i \rightarrow j$: record an **outlink** for page i
- If page $j \rightarrow i$: record an **inlink** for page i
- A page is not supposed to link to itself

Example: 4 web pages



4 \times 4 **adjacency matrix** C :

- *Outlink* for page $i \Rightarrow c_{j,i} = 1$
- Else $c_{j,i} = 0$

$$C = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Application: Google Eigenvector

Ranking r_i of page i determined by C

Example rules:

- 1 r_i should grow with the number of page i 's inlinks
- 2 r_i should be weighted by the ranking of each of page i 's inlinks
- 3 Let page i have an inlink from page j
then the more outlinks page j has, the less it should contribute to r_i

Not realistic but assume each page has at least one outlink and inlink

o_i : total number of outlinks of page i

Scale every element of column i by $1/o_i$

Google matrix D

$$d_{j,i} = \frac{c_{j,i}}{o_i}$$

Stochastic matrix: columns have non-negative entries and sum to one

Application: Google Eigenvector

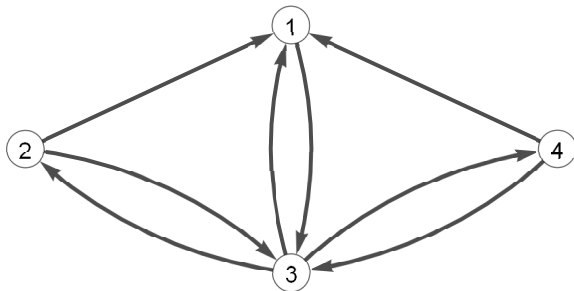
Connectivity/Adjacency matrix

$$C = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

\Rightarrow

Stochastic Google matrix

$$D = \begin{bmatrix} 0 & 1/2 & 1/3 & 1/2 \\ 0 & 0 & 1/3 & 0 \\ 1 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/3 & 0 \end{bmatrix}$$



Application: Google Eigenvector

Finding r_i involves knowing the ranking of all pages including r_i !

— Seems like an ill-posed circular problem, but ...

Find $\mathbf{r} = D\mathbf{r}$ where $\mathbf{r} = [r_1, \dots, r_N]^T$

— Eigenvector of D corresponding to the eigenvalue 1

— All stochastic matrices have an eigenvalue 1

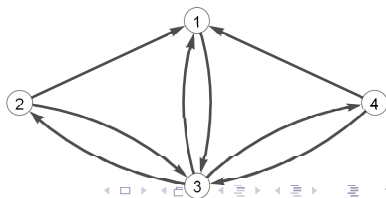
— \mathbf{r} is called a **stationary vector**

— 1 is D 's largest (dominant) eigenvalue

— Employ the *power method*

— Vector \mathbf{r} now contains the **page rank**

$\mathbf{r} = [0.67, 0.33, 1, 0.33]^T$
⇒ Highest ranked: page 3



Application: Google Eigenvector

In the real world — in 2021 — approximately 3.5 billion webpages
⇒ World's largest matrix to be used

Luckily it contains mostly zeroes — *sparse matrix*

Introduction Figure illustrates a Google matrix for ≈ 3 million pages

In the real world many more rules are needed and much more robust numerical analysis methods required

QR Algorithm

Triangular matrices prove to be helpful in developing efficient algorithms
— Continue that theme here with the QR algorithm

QR decomposition: $A = QR$

R : upper triangular matrix Q : orthogonal matrix

Utilize QR decomposition to solve eigenvalue problem $A\mathbf{u} = \lambda\mathbf{u}$

Orthogonal matrix Q suitable for a similarity transformation

$$A' = Q^T A Q \quad \Rightarrow \quad A, A' \text{ share same eigenvalues}$$

Write similarity transformation as

$$A' = RQ \quad \text{where } R = Q^T A$$

This process is repeated in the QR algorithm

QR Algorithm

Algorithm:

Let $A^{(1)} = A$

For $k = 1, 2, \dots$

Form the QR decomposition $A^{(k)} = Q^{(k)}R^{(k)}$

Set $A^{(k+1)} = R^{(k)}Q^{(k)}$

Elegantly simple algorithm!

If eigenvalues are distinct,

algorithm will return an approximation of the eigenvalues in $A^{(k)}$

Matrix $A^{(k)}$ will converge to an upper triangular matrix

— Eigenvalues on the diagonal

— If A is symmetric, then the matrix will be diagonal

Example:

$$A^{(1)} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \quad \lambda_i \approx 3.4, 2, 0.58$$

Sample iterations:

$$A^{(2)} = \begin{bmatrix} 2.8 & -0.74 & 0 \\ -0.74 & 2.34 & 0.63 \\ 0 & 0.63 & 0.85 \end{bmatrix} \quad A^{(6)} = \begin{bmatrix} 3.4 & -0.13 & 0 \\ -0.13 & 2.0 & 0.004 \\ 0 & 0.004 & 0.58 \end{bmatrix}$$

More iterations results in $A^{(k)}$ becoming closer to diagonal and improvements on the exact eigenvalues are made.

Eigenfunctions

Explore the space of all real-valued functions — **function space**

Do eigenvalues and eigenvectors have meaning there?

Let f be a function: $y = f(x)$ where x and y are real numbers

— Assume that f is smooth or differentiable

— Example: $f(x) = \sin(x)$

— The set of all such functions f forms a linear space

Define linear maps for elements of this function space

— Example: $Lf = 2f$

— Example: Derivatives $Df = f'$

To any function f the map D assigns another function

Example: let $f(x) = \sin(x)$ then $Df(x) = \cos(x)$

Eigenfunctions

How can we marry the concept of eigenvalues and linear maps?

D will not have *eigenvectors* since our linear space consists of functions

Instead: *eigenfunctions*

A function f is an eigenfunction of linear map D if

$$Df = \lambda f$$

D may have many eigenfunctions each corresponding to a different λ

Eigenfunctions

Any function f satisfying

$$f' = \lambda f$$

is an eigenfunction of the derivative map

The function $f(x) = e^x$ satisfies

$$f'(x) = e^x \quad \text{which may be written as} \quad Df = f = 1 \times f$$

$\Rightarrow 1$ is an eigenvalue of the derivative map D

More generally: all functions $f(x) = e^{\lambda x}$ satisfy (for $\lambda \neq 0$):

$$f'(x) = \lambda e^{\lambda x} \quad \text{which may be written as} \quad Df = \lambda f$$

\Rightarrow all real numbers $\lambda \neq 0$ are eigenvalues of D

Corresponding eigenfunctions are $e^{\lambda x} \Rightarrow$ infinitely many eigenfunctions!

Eigenfunctions

Example: second derivative $Lf = f''$

A set of eigenfunctions for this map is $\cos(kx)$ for $k = 1, 2, \dots$

$$\frac{d^2 \cos(kx)}{dx^2} = -k \frac{d \sin(kx)}{dx} = -k^2 \cos(kx)$$

and the eigenvalues are $-k^2$

Eigenfunctions

Eigenfunctions have many uses

- Differential equations
- Mathematical physics
- Engineering mathematics:
orthogonal functions key for data fitting and vibration analysis

Orthogonal functions arise as solution to a Sturm-Liouville equation

$$y''(x) + \lambda y(x) = 0 \quad \text{such that} \quad y(0) = 0 \quad \text{and} \quad y(\pi) = 0$$

- Solution: $y(x) = \sin(ax)$ for $a = 1, 2, \dots$
 - These are eigenfunctions of the Sturm-Liouville equation
 - The corresponding eigenvalues are $\lambda = a^2$

See Section 15.8 Application: Influenza Modeling

- eigenvalue
- eigenvector
- characteristic polynomial
- eigenvalues and eigenvectors of a symmetric matrix
- dominant eigenvalue
- eigendecomposition
- trace
- quadratic form
- positive definite matrix
- power method
- QR algorithm
- max-norm
- connectivity matrix
- adjacency matrix
- directed graph
- stochastic matrix
- stationary vector
- orthogonally diagonalizable
- similarity transformation
- eigenfunction