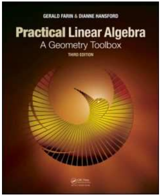
Practical Linear Algebra: A Geometry Toolbox

Third edition

Chapter 15: Eigen Things Revisited

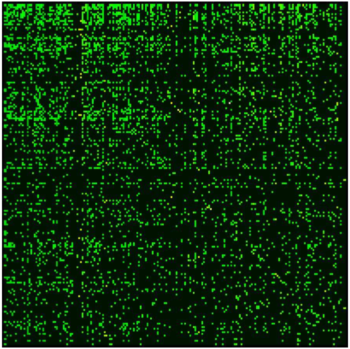
Gerald Farin & Dianne Hansford

CRC Press, Taylor & Francis Group, An A K Peters Book
www.farinhansford.com/books/pla



Outline

Connectivity matrix for a Google matrix



Chapter 7: 2×2 matrices
Here: $n \times n$ matrices

Eigenvalues and eigenvectors reveal action and geometry of map

Important in many areas:
— characterizing harmonics of musical instruments
— moderating movement of fuel in a ship
— analysis of large data sets
Google matrix:
Used to find the webpage ranking
(See Section: Google Eigenvector)

The Basics Revisited

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

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

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

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

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

Polynomial of degree n in λ — its zeroes are A's *eigenvalues*

The Basics Revisited

Example:

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

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

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

$$\lambda_1 = 4 \quad \lambda_2 = 3 \quad \lambda_3 = 2 \quad \lambda_4 = 1$$

Convention: order the eigenvalues in decreasing order

Dominant eigenvalue: largest eigenvalue in absolute value

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 $\lambda_2 = 1$

Instead: use diagonalization — see Chapter 16.

The Basics Revisited

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

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

Let $\lambda = 0$ then $p(0) = \det A = \lambda_1 \lambda_2 \cdot \ldots \cdot \lambda_n$

Finding zeroes of n^{th} degree polynomial nontrivial
— Use iterative method to find dominant eigenvalue (see next Section)
— Symmetric matrices always have real eigenvalues
— 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

Sometimes eigenvectors normalized to eliminate this ambiguity

Example: Find the eigenvectors

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

Starting with $\lambda_1 = 4$:

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

Repeating for all eigenvalues

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

Multiple zeroes of the characteristic polynomial
$\Rightarrow$ identical homogeneous systems $[A - \lambda I]\mathbf{r} = \mathbf{0}$
Example:

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

For $\lambda_1 = \lambda_2 = 2$

$$\begin{bmatrix} -1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{r}_1 = \mathbf{0} \qquad \Rightarrow \qquad \mathbf{r}_1 = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$$

For $\lambda_3 = 1$

$$\begin{bmatrix} 0 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{r}_3 = \mathbf{0} \qquad \Rightarrow \qquad \mathbf{r}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

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

The Basics Revisited

Symmetric matrix A:
— real eigenvalues
— eigenvectors are orthogonal

$\Rightarrow$ A is diagonalizable:
Possible to transform A to diagonal matrix $\Lambda = R^{-1}AR$
— Columns of R are A's eigenvectors
— Λ is a diagonal matrix of A's eigenvalues
— eigendecomposition of A

The Basics Revisited

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

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

Corresponding eigenvectors

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

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

The Basics Revisited

Projection matrices:

— eigenvalues are one or zero

0: eigenvector projected to the zero vector

$\Rightarrow$ determinant is zero and matrix is singular

1: eigenvector projected to itself

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

$\Rightarrow$ dimension is k and null space is dimension $n - k$

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

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

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

$$\lambda_{1,2} = 0 \quad \Rightarrow \quad \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 0 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix} \quad \Rightarrow \quad \mathbf{r}_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

Or find two eigenvectors that span 2D null space:

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

Trace of matrix A

$$\text{tr}(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n = a_{1,1} + a_{2,2} + \ldots + a_{n,n}$$

Gives insight to eigenvalues without computing directly

For 2×2 matrices

$$\det[A - \lambda I] = \lambda^2 - \lambda \text{tr}(A) + \det A \quad \Rightarrow \quad \lambda_{1,2} = \frac{\text{tr}(A) \pm \sqrt{\text{tr}(A)^2 - 4 \det A}}{2}$$

Example:

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

$\text{tr}(A) = -1$ and $\det A = -2 \quad \Rightarrow \quad \lambda_{1,2} = \frac{-1 \pm 3}{2}$

The Basics Revisited

Quadratic forms in $\mathbb{R}^n$

$$f(\mathbf{v}) = \mathbf{v}^{\mathrm{T}} C \mathbf{v} = c_{1,1} v_1^2 + 2c_{1,2} v_1 v_2 + \ldots + c_{n,n} v_n^2$$

The 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 $1/\sqrt{\lambda_n}$

Positive definite matrix: A real matrix satisfying

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

A: symmetric $n \times n$ matrix
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

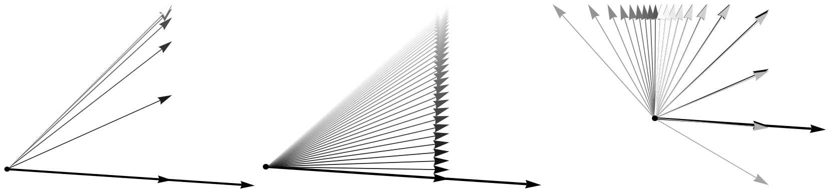
Start with arbitrary (nonzero) $\mathbf{r}^{(1)}$ — construct vector sequence

$$\mathbf{r}^{(i+1)} = A\mathbf{r}^{(i)}; \qquad i = 1, 2, \ldots$$

After a sufficiently large i $\mathbf{r}^{(i)}$ will begin to line up with $\mathbf{r}$: $\mathbf{r}^{(i+1)} = \lambda \mathbf{r}^{(i)}$

$\Rightarrow$ All components of $\mathbf{r}^{(i+1)}$ and $\mathbf{r}^{(i)}$ are (approximately) related by

$$\frac{r_j^{(i+1)}}{r_j^{(i)}} = \lambda \quad \text{for } j = 1, \ldots, n \qquad (*)$$



Longest black vector: initial guess; Successive iterations lighter shades

Each iteration scaled with respect to the ∞-norm

The Power Method

Algorithm:
Rather than checking each ratio (*) use the ∞-norm to define λ

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 $k = 2, \ldots, m$

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

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

The Power Method

Example: A_1, A_2, A_3 correspond to Figure from left to right

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

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

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

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

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

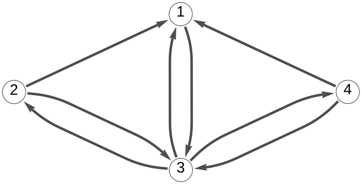
How can we rank all web pages?

Application: Google Eigenvector

Basics:

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

Example: 4 web pages



Directed graph

4×4 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

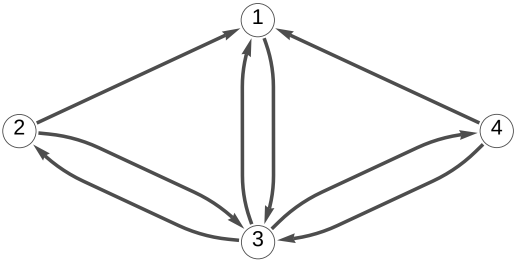
Application: Google Eigenvector

Adjacency matrix

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

Stochastic Google matrix

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

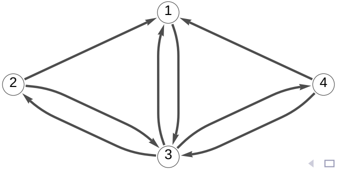


Application: Google Eigenvector

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 $\mathbf{r} = [r_1, \ldots, r_N]^{\mathrm{T}}$
— 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

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



Application: Google Eigenvector

In the real world — in 2013 — approximately 50 billion webpages
$\Rightarrow$ World's largest matrix to be used

Luckily it contains mostly zeroes — *sparse matrix*

Introduction Figure illustrates a Google matrix for $\approx$3 million pages

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

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 eigen*vectors* since our linear space consists of functions,
Instead: eigen*functions*
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

$$f' = \lambda f$$

Any function f satisfying this is an eigenfunction of the derivative map

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

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

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

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

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

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

Corresponding eigenfunctions are $e^{\lambda x}$

This map D has infinitely many eigenfunctions!

Eigenfunctions

Example: the map is the second derivative $Lf = f''$

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

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

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 result of the solution to a Sturm-Liouville equation

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

— Linear second order differential equation with boundary conditions
— Defines a *boundary value problem*
— Unknown are the functions $y(x)$ that satisfy this equation
— Solution: $y(x) = \sin(ax)$ for $a = 1, 2, \ldots$
— These are eigenfunctions of the Sturm-Liouville equation
— The corresponding eigenvalues are $\lambda = a^2$

WYSK

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