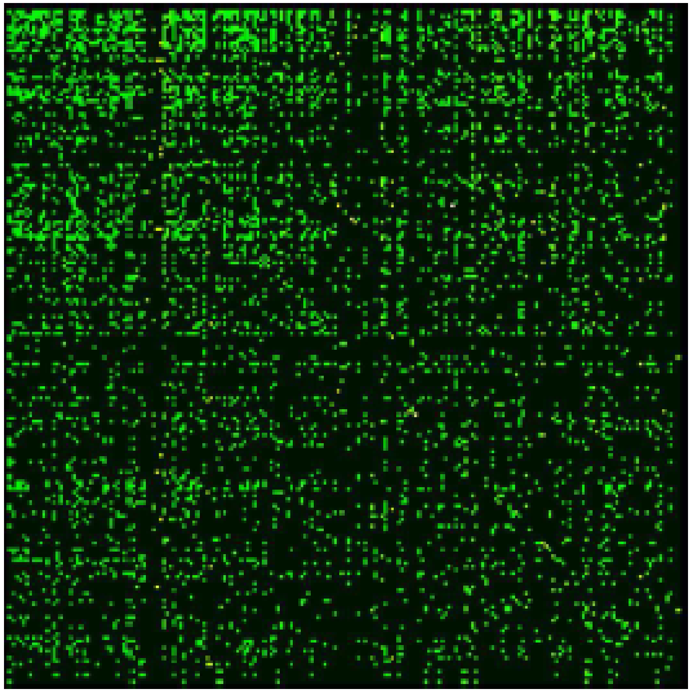
CS 357: Numerical Methods

Lecture 13: Eigenvalues and Eigenvectors

Eric Shaffer

Eigenthings



Eigenvectors and eigenvalues

Reveal action and geometry

Lots of applications

Engineering design
    make sure your bridge won't
    spontaneously fall down

Data analysis
    Google and PageRank

Eigenvectors

Given matrix A: which vectors $\mathbf{r}$ mapped to a multiple of itself?

$$A\mathbf{r} = \lambda\mathbf{r} \qquad \lambda \in \mathbb{R}$$

Disregard the "trivial solution" $\mathbf{r} = \mathbf{0}$

In 2D: at most two directions
Symmetric matrices: directions orthogonal (more on that later)

Fixed directions called the eigenvectors
— from the German word "*eigen*" meaning special or proper
Factor λ called its eigenvalue
Key to understanding geometry of a matrix

Finding Eigenvalues by Hand

How to find the eigenvalues of a 2×2 matrix A

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

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

Matrix $[A - \lambda I]$ maps a nonzero vector $\mathbf{r}$ to the zero vector
$\Rightarrow [A - \lambda I]$ rank deficient matrix $\Rightarrow$

$$p(\lambda) = \det[A - \lambda I] = 0$$

Characteristic equation: polynomial equation in λ
— 2D: characteristic equation is quadratic
$p(\lambda)$ called the characteristic polynomial

Finding Eigenvalues by Hand

Example:

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

$$p(\lambda) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0$$

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

$$\lambda_1 = 3 \qquad \lambda_2 = 1$$

Recall quadratic equation:
$a\lambda^2 + b\lambda + c = 0$ has the solutions

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Finding Eigenvalues by Hand

Example continued

Find $\mathbf{r}_1$ and $\mathbf{r}_2$ corresponding to $\lambda_1 = 3$ and $\lambda_2 = 1$

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

Homogeneous system and rank 1 matrix

$\Rightarrow$ infinitely many solutions

Forward elimination results in

$$\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{r}_1 = \mathbf{0}$$

Assign $r_{2,1} = 1$, then $\mathbf{r}_1 = c \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

Finding Eigenvalues by Hand

Next: $\lambda_2 = 1$, find $\mathbf{r}_2$

$$\begin{bmatrix} 2-1 & 1 \\ 1 & 2-1 \end{bmatrix} \mathbf{r}_2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{r}_2 = \mathbf{0}$$

$$\mathbf{r}_2 = c \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Recheck Figure: $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is not stretched – it is mapped to itself

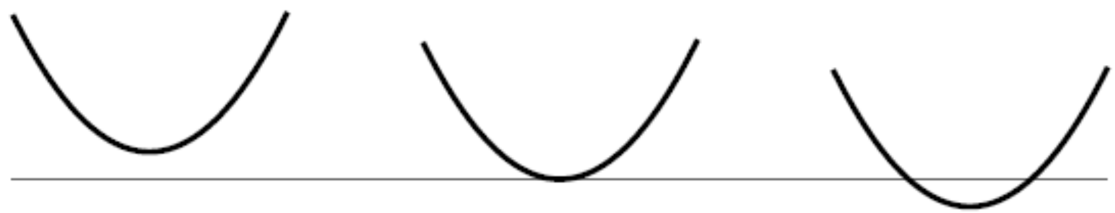
Often eigenvectors normalized for degree of uniqueness

$$\mathbf{r}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{r}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Dominant eigenvector: eigenvector corresponding to dominant eigenvalue

The Geometry of Eigenvectors

Quadratic polynomials have either no, one, or two *real* zeroes



If there are no zeroes: then A has no fixed directions

Example: rotations — rotate every vector; no direction unchanged

Rotation by -90°

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Characteristic equation

$$\begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = 0 \quad \Rightarrow \quad \lambda^2 + 1 = 0$$

Three Ways to Say the Same Thing

$Ax=\lambda x$

$(A-\lambda I)x=0$

$x \in N(A-\lambda I)$

Matrices with Easily Found Eigenthings

- Triangular Matrices
  - Eigenvalues are the diagonal entries
  - To find eigenvectors, find $N(A-\lambda I)$

Example

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

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

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

Finding Eigenvalues for Larger n

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 $1/\lambda_i$

Eigenvalues and Changes to A

- Scaling $\beta A x = \beta \lambda x$
- Shift $(A - \sigma I)x = Ax - \sigma \mathrm{x} = (\lambda - \sigma)\mathrm{x}$
- Matrix Power $A^k x = \lambda^k x$
- Inverse $A^{-1} x = \frac{1}{\lambda} x$

Similarity Transform

- Previous transforms left eigenvectors same, changed eigenvalues
- Can we change eigenvectors and retain eigenvalues?
- Let T be an invertible matrix. A *similarity transform* is

$$T^{-1}AT$$

- Let $y = T^{-1}x$

then $(T^{-1}AT)y = T^{-1}ATT^{-1}x = T^{-1}Ax = \lambda T^{-1}x = \lambda y$

Does every Square Matrix have n Linearly Independent Eigenvectors

- If A has n linearly independent eigenvectors xi we can create matrix X such that

$$X = [x1 \quad \ldots \quad xn]$$

$$AX = X\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = \mathrm{XD}$$

$$X^{-1}AX = D$$

- A is called ***diagonalizable***
- ***Not all matrices are diagonalizable***

Rayleigh Quotient

- We can estimate an eigenvalue if we know an eigenvector x
- Rayleigh quotient $$\frac{x \cdot Ax}{x \cdot x} = \frac{x \cdot \lambda x}{x \cdot x} = \lambda$$

The Power Method

- A is an n x n diagonalizable matrix, x an eigenavector
- It follows that $A^i x = \lambda^i x$
- This can be used to find the eigenvector x

The Power Method

Given A, and n x n diagonalizable matrix
Pick an arbitrary vector x_1

$x_k = Ax_{k-1}$

Keep going until the ratio $\dfrac{\|x_k\|}{\|x_{k-1}\|}$ stops changing much

Why does this work?

Why does this work?

The Power Method

What does $\dfrac{\|x_k\|}{\|x_{k-1}\|}$ converge to?

The Power Method

- If λ is large or small there could be numerical problems
  - What kinds?

- Can it converge if first guess is perpendicular to the eigenvector?
  - Yes…round off error means it will converge slowly…

- If $|\lambda_1| \approx |\lambda_2|$ method may not converge
- Method limited to symmetric matrices with a dominant eigenvalue

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

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

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

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

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.

Normalized Power Iteration

What does this converge to?

$$x_k = \frac{Ax_{k-1}}{||Ax_{k-1}||}$$

Why would anyone do that?

Inverse Iteration

$$x_k = \frac{A^{-1}x_{k-1}}{\|A^{-1}x_{k-1}\|}$$

What does this converge to?

Why would anyone do that?

Inverse Iteration with Shift

What does this converge to?

$$x_k = \frac{(A - \sigma I)^{-1} x_{k-1}}{\|(A - \sigma I)^{-1} x_{k-1}\|}$$

Why would anyone do that?

Rayleigh Quotient Iteration

$$x_k = \frac{(A - \sigma I)^{-1} x_{k-1}}{\|(A - \sigma I)^{-1} x_{k-1}\|}$$