Eigenvalue problems
Iterative methods

Definition (Eigenvalue and Eigenvector)
An eigenvector $x \neq 0$ of a matrix $A \in \mathbb{R}^{n \times n}$ is any vector satisfying $Ax = \lambda x$ for $\lambda \in \mathbb{R}$; the corresponding λ is called an eigenvalue

$\Rightarrow (A - \lambda I)x = 0$ has a non-trivial solution
This will be true if and only if
$$\det(A - \lambda I) = 0$$
$$\Downarrow$$
The resulting polynomial is called the characteristic polynomial.

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

$A - \lambda I = \begin{bmatrix} 3-\lambda & 2 \\ 1 & 4-\lambda \end{bmatrix}$

$\det(A - \lambda I) = 0 \Rightarrow (3-\lambda)(4-\lambda) - 2 = 0$
$12 - 7\lambda + \lambda^2 - 2 = 0$
$\underline{\lambda^2 - 7\lambda + 10 = 0}$

characteristic polynomial

$\Rightarrow (\lambda - 2)(\lambda - 5) = 0$
The eigenvalues are $\lambda_1 = 5$ and $\lambda_2 = 2$

Definition The spectrum of A is the set of eigenvalues of A. The spectral radius $\rho(A)$ is the maximum value $|\lambda|$ over all eigenvalues of A.

Example $A = \begin{bmatrix} 0 & -4 \\ -0.5 & 1 \end{bmatrix}$

$A - \lambda I = \begin{bmatrix} -\lambda & -4-0 \\ -0.5-0 & 1-\lambda \end{bmatrix} = \begin{bmatrix} -\lambda & -4 \\ -1/2 & 1-\lambda \end{bmatrix}$

$\det(A - \lambda I) = 0 \Rightarrow -\lambda(1-\lambda) - 2 = 0$
$\lambda^2 - \lambda - 2 = 0 \Rightarrow (\lambda + 1)(\lambda - 2) = 0$

spectrum of $A = \{-1, 2\}$
$\rho(A) = \underline{\underline{2}}$

Proposition Every matrix $A \in \mathbb{R}^{n \times n}$ has at least one (potentially complex) eigenvector.

proof $x \in \mathbb{R}^n \setminus 0$; x is not zero vector
Also $A \neq 0$ (Exclude zero matrix)
we study the following set:
$\{x, Ax, A^2 x, \ldots, A^n x\}$
$\Rightarrow$ linearly dependent
($n+1$ vectors in $\mathbb{R}^n$)
Then, it follows that there exists a non-trivial combination to form 0.

$$c_0 x + c_1 Ax + \ldots + c_n A^n x = 0$$

$\Downarrow$

This motivates us to define the following polynomial

$$f(z) \equiv c_0 + c_1 z + \ldots + c_n z^n$$

From fundamental theorem of algebra there exist $m \geq 1$ roots $z_i \in \mathbb{C}$ and $c \neq 0$ such that

$$f(z) = c(z - z_1)(z - z_2) \ldots (z - z_m)$$

$$c_0 x + c_1 Ax + \ldots c_n A^n x = 0$$

$$(c_0 I_{n \times n} + c_1 A + \ldots + c_n A^n) x = 0$$

$\downarrow$

(Entry wise this is $(n+1)$ degree polynomial)

$$\Rightarrow c(A - z_1 I_{n \times n}) \cdots (A - z_m I_{n \times n}) = 0$$

claim At least one $A - z_i I_{n \times n}$ has a null space
(otherwise each term would be invertible $\Rightarrow x = 0$)

Let v be the non-zero vector in the null space of $A - z_i I_{n \times n}$

$$\Rightarrow \quad Av = z_i v$$

Proposition Eigenvectors corresponding to different eigenvalues must be linearly independent

Proof Proof by induction
$n = 1$ is trivial. (why?)
Suppose that the statement is true for $r-1$. Assume that there is a non-trivial linear combination

$$c_1 v_1 + c_2 v_2 + \ldots + c_r v_r = \sum_{k=1}^{r} c_k v_k = 0 \quad *$$

($v_1, v_2, \ldots, v_r$ are eigenvectors)

Apply $(A - \lambda_r I)$ to $*$

$$(A-\lambda_r I) v_1 = A v_1 - \lambda_r v_1$$
$$= \lambda_1 v_1 - \lambda_r v_1$$
$$= [\lambda_1 - \lambda_r] v_1$$

Also note that

$$(A - \lambda_r I) v_r = A v_r - \lambda_r v_r$$
$$= \lambda_r v_r - \lambda_r v_r$$
$$= 0$$

With this, we obtain

$$\sum_{k=1}^{r-1} c_k (\lambda_k - \lambda_r) v_k = 0$$

By induction, $v_1 \dots v_{r-1}$ are linearly independent

$\Downarrow$

$c_k(\lambda_k - \lambda_r) = 0$ (Since $\lambda_k \neq \lambda_r$)

$\Rightarrow c_k = 0$ for $k < r$

Now consider $*$

$$\sum_{k=1}^{r} c_k v_k = 0 \Rightarrow c_r v_r = 0 \Rightarrow c_r = 0$$

$\therefore$ Trivial linear combination

$(v_1 \dots v_r)$ are linearly independent

Remark • An $n \times n$ matrix can have at most n distinct eigenvalues

A matrix is nondefective or diagonalizable if its eigenvectors span $\mathbb{R}^n$

Example $A = \begin{pmatrix} 5 & 2 \\ 0 & 5 \end{pmatrix}$ $\lambda_1 = 5$ $\lambda_2 = 5$

$$(A - \lambda I) x = 0 \Rightarrow \begin{pmatrix} 5-\lambda & 2 \\ 0 & 5-\lambda \end{pmatrix} \vec{x} = \vec{0}$$

characteristic polynomial : $(5-\lambda)^2 \cancel{\vec{x}} = 0$ $\lambda = 5$

$$\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Longrightarrow \begin{matrix} x_2 = 0 \\ x_1 = \text{free} \end{matrix}$$

Eigenvector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Non-diagonalizable or defective

Diagonalizability Let the eigenvectors $u_1 \dots u_n$ be linearly independent

$$\Rightarrow \quad AU_1 = \lambda_1 U_1$$
$$AU_2 = \lambda_2 U_2$$
$$\vdots$$
$$AU_n = \lambda_n U_n$$

Equivalent to $AU = UD$

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

$$A = UDU^{-1}$$

A and D are similar

Proposition All eigenvalues of a symmetric matrix are real. Eigenvectors corresponding to different eigenvalues are orthogonal.

Proof Exercise! or read any standard text on linear algebra

Proposition Suppose A is a symmetric matrix. Then A has exactly n orthonormal eigenvectors.

$$A = UDU^T \qquad \text{[spectral theorem]}$$

Theorem $\det(A) = \prod_{j=1}^{n} \lambda_j \qquad \text{Trace}(A) = \sum_{j=1}^{n} \lambda_j$

proof Exercise! or refer to a standard linear algebra text

Why not directly solve the characteristic polynomial?

Theorem For any $n \geq 5$, there is a polynomial $p(z)$ of degree n with rational coefficients that has a real root $p(r) = 0$ with the property that r cannot be written using any expression involving rational numbers, addition, substraction, division and k^{th} roots

Main Idea Even in exact arithmetic, exact roots could not be produced using finite number of steps.

$\therefore$ Eigenvalue solvers must be iterative!

Power iteration

Definition An eigenvalue λ of A is dominating if its absolute value is strictly greater than the absolute values of all other eigenvalues

$\Rightarrow$ The corresponding eigenvector is called a dominating eigenvector

Example $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

$\lambda_1 = 1 \qquad \lambda_2 = -1$

No dominant eigenvalue

Example $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ No dominant eigenvalue

$\lambda_1 = 2 \qquad \lambda_2 = 2 \qquad \lambda_3 = 1$

Example $A = \begin{bmatrix} 0 & -4 \\ -0.5 & 1 \end{bmatrix}$

$\lambda_1 = 2 \qquad \lambda_2 = -1$

Dominant eigenvalue 2

Assume (for ease) A is diagonalizable

$$AU_i = \lambda_i U_i \qquad i = 1, 2, \dots n$$

$U_1, U_2, \dots U_n$ are linearly independent

$x \in \mathbb{R}^n$ is some vector

$$x = \sum_{i=1}^{n} c_i U_i$$

$$x = \sum_{i=1}^{q} c_i U_i + \sum_{q+1}^{n} c_i U_i$$

$U_1 \dots U_q$ correspond to $\lambda_1 = \lambda$

$$Ax = \lambda\left(\sum_{i=1}^{q} c_i U_i\right) + \sum_{q+1}^{n} c_i \lambda_i U_i$$

$$A^k x = \underbrace{\lambda^k\left(\sum_{i=1}^{q} c_i U_i\right)}_{1} + \underbrace{\sum_{q+1}^{n} c_i \lambda_i^k U_i}_{2}$$

Compare the sizes of ① and ②

$\lambda^k \left(\sum_{i=1}^{q} c_i u_i \right)$ for greater than $\sum_{q+1}^{n} c_i \lambda_i^k u_i$

in absolute value.

* As long as $\sum_{i=1}^{q} c_i u_i \neq 0$

$$A^k x = \lambda^k \left[\left(\sum_{i=1}^{q} c_i u_i \right) + \sum_{q+1}^{n} c_i \left(\frac{\lambda_i^k}{\lambda^k} \right) u_i \right]$$

As $k \to \infty$ $\quad \to 0$

$\Downarrow$

points in direction of dominating eigenvector

<u>Power iteration for computing dominating real eigenvectors</u>

Let $A \in \mathbb{R}^{n \times n}$ be a matrix that has a dominating eigenvalue $\lambda \in \mathbb{R}$. Choose a random $x \in \mathbb{R}^m$ and compute the sequence

$$x_0, x_1, x_2, \ldots$$

defined recursively as

$$z_{k+1} = A x_k \qquad x_{k+1} = \frac{z_{k+1}}{\|z_{k+1}\|_2} \qquad k = 0, 1, 2, \ldots$$

$\lambda > 0 \quad \lim_{k \to \infty} x_k =$ dominating eigenvector of A

$\lambda < 0 \quad \lim_{k \to \infty} (-1)^k x_k =$ dominating eigenvector of A

<u>Questions</u> ① What if the initial guess is such that $c_1 = c_2 = \ldots = c_q = 0$?

$\Rightarrow$ Rare occurrence

$\Rightarrow$ Initialize randomly

$\Rightarrow$ floating point Error can be useful!

② What happens if λ and $-\lambda$ are the eigenvalues with the largest magnitudes?

$\Rightarrow$ It can fail

* Useful when A is sparse
* Power iteration could only find eigenvector corresponding to the dominant eigenvalue.