Eigenvalues let A be an $n\times n$ matrix and $x \in \mathbb{C}^n$ be nonzero.

If $\exists \lambda \in \mathbb{C}$ s.t. $Ax = \lambda x$ then x is called an eigenvector corresponding to the eigenvalue λ

Example let $A = \begin{pmatrix} 0 & 4 \\ -1 & 0 \end{pmatrix}$.

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 4 \\ -1 & -\lambda \end{vmatrix} = \underbrace{\lambda^2 + 4}_{\text{"characteristic polynomial"}} = 0$$

$\Rightarrow$ $\lambda = \pm 2i$ are the eigenvalues of A.

Note A is a real matrix but eigenvalues are complex (in general)

For small matrices, finding roots of $\det(A - \lambda I)$ by hand is useful but in general we may run into issues!

5.1 Matrix Eigenvalue Problem: Power Method

let $\lambda_1, \cdots \lambda_n$ be eigenvalues with associated eigenvectors $u^{(1)}, \ldots u^{(n)}$ s.t.

1) $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$
2) $\{u^{(1)}, \ldots, u^{(n)}\}$ is a basis for $\mathbb{C}^n$. (spanning set that is linearly independent)

let $x^{(0)} = a_1 u^{(1)} + a_2 u^{(2)} + \cdots + a_n u^{(n)}$ s.t. $a_1 \neq 0$

We form the sequence

$$x^{(1)} = A x^{(0)}, \quad x^{(2)} = A x^{(1)}, \cdots \qquad x^{(k+1)} = A x^{(k)}$$

or

$$x^{(k)} = A^k x^{(0)}$$

To simplify our analysis, we assume $x^{(0)} = u^{(1)} + u^{(2)} + \cdots + u^{(n)}$ without any less of generality.

$$x^{(k)} = A^k x^{(0)} = A^k u^{(1)} + \cdots + A^k u^{(n)}$$

$$x^{(k)} = \lambda_1^k u^{(1)} + \cdots + \lambda_n^k u^{(n)}$$

$$x^{(k)} = \lambda_1^k \Big(u^{(1)} + \underbrace{\left(\frac{\lambda_2}{\lambda_1}\right)^k u^{(2)} + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k u^{(n)}}_{\to 0 \text{ as } k\to\infty} \Big)$$

Since $|\lambda_j| < |\lambda_1|$ for $j \geq 2$, $\left(\frac{\lambda_j}{\lambda_1}\right)^k \to 0$ as $k \to \infty$.

Set $x^{(k)} = \lambda_1^k \left(u^{(1)} + \varepsilon^{(k)} \right)$

where $\varepsilon^{(k)} = \left(\frac{\lambda_2}{\lambda_1}\right)^k u^{(2)} + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)^k u^{(n)} \to 0$ as $k \to \infty$

let φ be any linear functional on $\mathbb{C}^n$, that is, $\varphi : \mathbb{C}^n \to \mathbb{C}$ and it is linear: $\varphi(x+y) = \varphi(x) + \varphi(y)$ and $\varphi(\lambda x) = \lambda \varphi(x)$.

Then e.g. $\varphi\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 3x_1 + x_2$

$$\varphi(x^{(k)}) = \lambda_1^k \big(\varphi(u^{(1)}) + \underbrace{\varphi(\varepsilon^{(k)})}_{\to 0} \big)$$

$$\Rightarrow r_k := \frac{\varphi(x^{(k+1)})}{\varphi(x^{(k)})} = \frac{\lambda_1^{\cancel{k+1}} \left(\varphi(u^{(1)}) + \varphi(\varepsilon^{(k+1)}) \right)}{\cancel{\lambda_1^k} \left(\varphi(u^{(1)}) + \varphi(\varepsilon^{(k)}) \right)} \to \lambda_1$$

Since the direction of $x^{(k)}$ is approaching to the direction of $u^{(1)}$ as $k \to \infty$, we can also obtain $u^{(1)}$ with this method. $x^{(k)} = \lambda_1^k (u^{(1)} + \underbrace{\varepsilon^{(k)}}_{\to 0})$

Pseudo code

```
input n, A, x, M
output 0, x
for k = 1 to M do
    y ← Ax
    r ← φ(y)/φ(x)
    x ← y
    output k, x, r
end do
```

$x^{(k)} \sim x$ $y \sim x^{(k+1)}$

for some linear functional φ.

In general it is better to replace the line $x \leftarrow y$ by $x \leftarrow y/\|y\|$

Example Use the power method on the following matrix and initial vector

$$A = \begin{pmatrix} 6 & 5 & -5 \\ 2 & 6 & -2 \\ 2 & 5 & -1 \end{pmatrix} \qquad x = \begin{pmatrix} -1 \\ \boxed{1} \\ 1 \end{pmatrix}$$

Note $\|y^{(0)}\|_\infty = 6$

$x^{(0)} = x$ $y^{(0)} = A x^{(0)} = \begin{pmatrix} -6 \\ \boxed{2} \\ 2 \end{pmatrix}$ $x^{(1)} = y^{(0)} / \|y^{(0)}\|_\infty = \begin{pmatrix} -1 \\ 1/3 \\ 1/3 \end{pmatrix}$. Say $\varphi\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2$

then $\varphi(y^{(0)}) / \varphi(x^{(0)}) = 2 = r_0$

k	0	1	2	3	4		6		28
$x_1^{(k)}$	-1	-1 *	-1	-1	-1				-1
$x_2^{(k)}$	1	0.333	-0.111	-0.407	-0.605	...		...	-0.998
$x_3^{(k)}$	1	0.333	-0.111	-0.407	-0.605				-0.998
r_{k-1}	X	2	-2	22	8.909		6.715		6.0001

$\lambda_1 = 6$ and $u^{(1)} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ //

Thm If λ is an eigenvalue of an invertible matrix A, then λ^{-1} is an eigenvalue of A^{-1}!

Pf Say $Ax = \lambda x$ for $x \neq 0$. Then,

$x = A^{-1}Ax = A^{-1}\lambda x = \lambda A^{-1}x \Rightarrow \lambda^{-1}x = A^{-1}x$.

Inverse Power Method.

Given A with eigenvalues $\lambda_1, \dots \lambda_n$ s.t.

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{n-1}| > |\lambda_n| > 0$$

smallest eigenvalue is unique.

Then A^{-1} exists and its eigenvalues are $\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_n^{-1}$ satisfying

$$|\lambda_n^{-1}| > |\lambda_{n-1}^{-1}| \geq |\lambda_{n-2}^{-1}| \geq \dots \geq |\lambda_1| > 0$$

In principal we can apply the Power Method on A^{-1} to determine λ_n. However, instead of computing A^{-1} and $x^{(k+1)} = A^{-1}x^{(k)}$, we compute $Ax^{(k+1)} = x^{(k)}$ using Gaussian-elimination. (We carry out factorization phase only once!)

This is the inverse power method

Example

$$A = \begin{pmatrix} 6 & 5 & -5 \\ 2 & 6 & -2 \\ 2 & 5 & -1 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1/3 & 1 & 0 \\ 1/3 & \frac{10}{3} & 1 \end{pmatrix}}_{L} \underbrace{\begin{pmatrix} 6 & 5 & -5 \\ 0 & 13/3 & -1/3 \\ 0 & 0 & \frac{12}{13} \end{pmatrix}}_{U}$$

Say $x^{(0)} = \begin{pmatrix} 3 \\ 7 \\ -13 \end{pmatrix}$. Then we find $y^{(0)}$ by solving $Ay^{(0)} = x^{(0)}$

or equivalently $LUy^{(0)} = x^{(0)}$ or $Uy^{(0)} = L^{-1}x^{(0)}$

Then we set $x^{(1)} = y^{(0)}/\|y^{(0)}\|_\infty$ and continue similarly.

Using $\varphi\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1$

k	0	1	2		11	12
$x^{(k)}$	3 7 -13	-0.802 -0.008 -1.000	-0.951 -0.018 -1.000	---	-1.000 0.000 -1.000	-1.000 0.000 -1.000
r_{k-1}	X	-5.889	1.198		1.000	1.000

$r_k \to \lambda_3^{-1}$

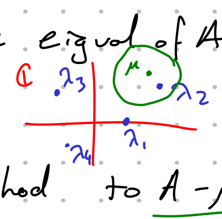
2

Thus, the smallest eigenvalue of A, λ_3, appears to satisfy

$\lambda_3^{-1} = 1$ i.e. $\lambda_3 = 1$ $\frac{1}{2}$ and corr. eig.vec is $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$.

Consider the shifted matrix $A - \mu I$. Assume μ is not an eigenval of A and there is a unique eigval of A, λ_k, s.t.

$0 < |\lambda_k - \mu| < \varepsilon$ for some $\varepsilon > 0$.



Then applying the inverse power method to $A - \mu I$ we can obtain λ_k. Exercise: why? (Hint: think about eig.values of $A - \mu I$)

(First, $r_* \to z = (\lambda_k - \mu)^{-1}$, then $\lambda_k = z^{-1} + \mu$)

Similarly if there is a unique λ_k s.t. $|\lambda_k - \mu| > \varepsilon > 0$ then the power method gives us λ_k.

Summary

Method	Equation	Computes
Power	$x^{(k+1)} = Ax^{(k)}$	largest eig.val.
inverse power	$Ax^{(k+1)} = x^{(k)}$	smallest eig.val.
shifted power	$x^{(k+1)} = (A - \mu I)x^{(k)}$	eig.val farthest from μ
shifted inverse power	$(A - \mu I)x^{(k+1)} = x^{(k)}$	eig.val closest to μ

5.3 Orthogonal Decompositions and Least Squares Problems

Basic Concepts

Recall $u, v \in \mathbb{C}^n$ are orthogonal if $\langle u, v \rangle = 0$

More generally, $\{v_1, v_2, \dots, v_k\}$ is orthogonal if $\langle v_i, v_j \rangle = 0$ for $i \neq j$.

It is orthonormal if $\langle v_i, v_j \rangle = \delta_{ij}$

$\Rightarrow \|v_i\|_2^2 = \langle v_i, v_i \rangle = \delta_{ii} = 1$

Recall $\langle x, y \rangle = y^* x$

$\langle v_i, v_j \rangle = \delta_{ij}$ $\delta_{ij} = (I)_{ij}$

Set $A_{n \times k} = [v_1 \; v_2 \; \cdots v_k]$.

$A^*A = I \overset{\text{Exercise}}{\iff} \{v_1 \cdots v_k\}$ is o.n. (orthonormal)

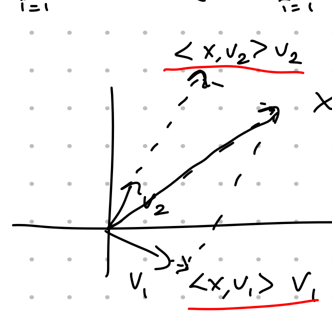
Say $[v_1 \cdots v_n]$ is a basis for $\mathbb{C}^n$. Then for $x \in \mathbb{C}^n$,

$\exists!$ "unique" $c_i \in \mathbb{C}$ s.t. $x = \sum_{i=1}^{n} c_i v_i$

$i \neq j$ $\delta_{ij} = 0$

$\Rightarrow \langle x, v_j \rangle = \langle \sum_{i=1}^{n} c_i v_i, v_j \rangle = \sum_{i=1}^{n} c_i \langle v_i, v_j \rangle = \sum_{i=1}^{n} c_i \delta_{ij} = c_j$

Thus, $x = \sum_{i=1}^{n} \langle x, v_i \rangle v_i$



$\{v_1, v_2\}$ are o.n.

Gram-Schmidt Process

Given a linearly indep't set $\{v_1, v_2, \dots v_k\}$ of vectors in $\mathbb{C}^n$ $(k \leq n)$ we can generate an o.n. set $\{u_1, \dots, u_k\}$ by

$$u_k = \frac{\left(v_k - \sum_{i<k} \langle v_k, u_i \rangle u_i\right)}{\left\| v_k - \sum_{i<k} \langle v_k, u_i \rangle u_i \right\|_2}$$

Exercise Prove that $U^*U = I$ for $U = [u_1 \cdots u_k]$

e.g. $u_1 = \dfrac{v_1 - \sum_{i<1} \langle v_1, u_i \rangle u_i}{\| v_1 - \sum_{i<1} \langle v_1, u_i \rangle u_i \|_2}$ Since $i<1$, $\sum_{i<1} = 0$

$u_1 = \dfrac{v_1}{\|v_1\|_2}$ unit vector.

$$u_2 = \frac{v_2 - \sum_{i<2} \langle v_2, u_i \rangle u_i}{\| v_2 - \sum_{i<k} \langle v_2, u_i \rangle u_i \|_2} = \frac{v_2 - \langle v_2, u_1 \rangle u_1}{\| v_2 - \langle v_2, u_1 \rangle u_1 \|_2}$$

$$\begin{aligned}
\langle u_1, u_2 \rangle &= C \langle v_1, v_2 - \langle v_2, u_1 \rangle u_1 \rangle \\
&= C\left(\langle v_1, v_2 \rangle - \langle v_1, \underline{\langle v_2, u_1 \rangle} u_1 \rangle \right. \\
&= C\left(\langle v_1, v_2 \rangle - \overline{\langle v_2, u_1 \rangle} \langle v_1, u_1 \rangle \right)
\end{aligned}$$

$u_1 = \dfrac{v_1}{\|v_1\|_2}$

$$\begin{aligned}
&= C\left(\langle v_1, v_2 \rangle - \overline{\left\langle v_2, \frac{v_1}{\|v_1\|_2} \right\rangle} \left\langle v_1, \frac{v_1}{\|v_1\|_2} \right\rangle \right) \\
&= C\left(\langle v_1, v_2 \rangle - \frac{1}{\cancel{\|v_1\|_2^2}} \overline{\langle v_2, v_1 \rangle} \cancel{\langle v_1, v_1 \rangle} \right) \\
&= C\left(\langle v_1, v_2 \rangle - \overline{\langle v_2, v_1 \rangle} \right) = C\left(\langle v_1, v_2 \rangle - \langle v_1, v_2 \rangle \right) \\
&= 0
\end{aligned}$$

Thus, $\langle u_1, u_2 \rangle = 0$