__/__/__

★ $\boxed{\text{Power iteration}}$ (assumption $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$)

Last time . • algorithm

• justification in the case A is diagonal.

Now, Suppose A is diagnolizable matrix i.e. $\exists$ an invertible matrix P

s.t. ~~$A = PAP^{-1}$~~

$A = P \Lambda P^{-1}$

~~& Λ~~

$\Lambda = \text{diag}(\lambda_1 \cdots \lambda_n)$ containing e-values of A.

Let $x_0 = P \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$, s.t $\xi_1 \neq 0$

$A = P \Lambda P^{-1}$

$A^i = P \Lambda^i P^{-1}$

$$A^i x_0 = P \Lambda^i P^{-1} . P \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$$

$$= P \Lambda^i \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = P \begin{pmatrix} \lambda_1^i \xi_1 \\ \vdots \\ \lambda_n^i \xi_n \end{pmatrix}$$

$$= \xi_1 \lambda_1^i P \begin{pmatrix} 1 \\ \left(\frac{\lambda_2}{\lambda_1}\right)^i \left(\frac{\xi_2}{\xi_1}\right) \\ \vdots \\ \left(\frac{\lambda_n}{\lambda_1}\right)^i \left(\frac{\xi_n}{\xi_1}\right) \end{pmatrix}$$

So

$$x_i = \frac{A^i x_0}{\|A^i x_0\|} = P \begin{pmatrix} 1 \\ \left(\frac{\lambda_2}{\lambda_1}\right)^i \frac{\varepsilon_2}{\varepsilon_1} \\ \vdots \\ \left(\frac{\lambda_n}{\lambda_1}\right)^i \frac{\varepsilon_n}{\varepsilon_1} \end{pmatrix}$$

$\rightarrow Pe_1$ as $i \rightarrow \infty$

$= P_1$ (1st element of P)

= e-vector corresponding to λ_1.

Remark: ① If A has a pair of complex conjugate as it's largest e-values (or even $k, -k \in \mathbb{R}$) then the power iteration will not proceed as expected.

② The rate of convergence of the power iteration depends on the gap b/t^n λ_1 & λ_2 ("eigen gap")

(if $\lambda_2 << \lambda$ then $\frac{\lambda_2}{\lambda_1} << 1$, so conv. is faster)

★ $\boxed{\text{Inverse Iteration}}$

Useful when we are given a good approx (or a value say σ, in the ~~unethodd~~ neighbourhood) of an e-value λ (unknown) & we want to find λ & it's corresponding e-vector

Idea: to apply the power iteration to a shifted matrix

__/__/___

Observation

① If $A = P \Lambda P^{-1}$

$$A - \sigma I = P \Lambda P^{-1} - P \sigma I P^{-1}$$

$$= P(\Lambda - \sigma I) P^{-1}$$

($\sigma I = \sigma I P P^{-1} = P \sigma I P^{-1}$
scalar matrices commute)

&

$$(A - \sigma I)^{-1} = P (\Lambda - \sigma I)^{-1} P^{-1}$$

$(A - \sigma I)^{-1}$ has the same e-vectors ~~e-value~~ as A & the corresponding e-values $(\lambda_K - \sigma)^{-1}$

② The largest e-value of $(A - \sigma I)^{-1}$ gives the e-value of A that is closest to σ.

($\frac{1}{\lambda_k - \sigma}$ is largest among $\left\{ \frac{1}{\lambda_i - \sigma} \right\}$

if $|\lambda_k - \sigma| < |\lambda_i - \sigma| \quad \forall \; i \neq k$

i.e. λ_K is closest to σ among all λ_i

Algorithm

x_0: initial vector

$i = 0$ to convergence

$$y_{i+1} = (A - \sigma I)^{-1} x_i$$

$$x_{i+1} = y_{i+1} / \|y_{i+1}\|_2$$

$$\lambda_{i+1} = x_{i+1}^T A x_{i+1}$$

end.

$$(A-\sigma I)^{-i}x_0 = P\,(A-\sigma I)^{-i}\,P^{-1}P \overbrace{\begin{bmatrix}\varepsilon_1\\ \vdots\\ \varepsilon_n\end{bmatrix}}^{x_0}$$

$$= P\begin{pmatrix}\varepsilon_1\left(\frac{1}{\lambda_1-\sigma}\right)^i\\ \varepsilon_2\left(\frac{1}{\lambda_2-\sigma}\right)^i\\ \vdots\\ \varepsilon_n\left(\frac{1}{\lambda_n-\sigma}\right)^i\end{pmatrix}$$

$$= \varepsilon_k\,(\lambda_k-\sigma)^{-i}\,P\begin{pmatrix}\frac{\varepsilon_1}{\varepsilon_k}\left(\frac{\lambda_k-\sigma}{\lambda_1-\sigma}\right)^i\\ \vdots\\ \vdots\end{pmatrix}$$

$$x_i = \frac{(A-\sigma I)^{-i}x_0}{\|\cdots\|}$$

$$= P\begin{pmatrix}\\ \\ \end{pmatrix}$$

$$\to Pe_k \quad \text{as } i\to\infty.$$

$$= p_k$$

III) $\boxed{\text{Block power method}}$ | orthogonal iteration. / Simultaneous iteration. / Subspace iteration.

The idea is to apply power iteration to several vectors at once.

__/__/____

So, instead of starting with one vector x_0 we start with an orthogonal ~~a~~ set of ~~linearly independent~~ initial vectors ~~a set of~~

$\{x_1^{(0)} \dots x_n^{(0)}\} \rightarrow$ assumption e-values satisfy

$\lambda_1 > \lambda_2 \dots \lambda_n > \lambda_{n+1}$.

let $A \in \mathbb{C}^{m \times m}$ Under suitable assumptions it can be shown that $\langle A^i x_1^{(0)} \dots , A^i x_n^{(0)} \rangle$ converges to the span of the e-vectors of A corresponding to the n Largest e-values of A.

largest $\equiv$ largest in abs. value

let $z_0 = \left[x_1^{(0)} \middle| \dots \; x_n^{(0)} \right]$

let Z_K denote

$\left[A^K x_1^{(0)} \middle| \dots \middle| A^K x_n^{(0)} \right]$

Algorithm

Z_0 be an $m \times n$ orthogonal matrix

$i = 0$ to convergence.

$Z_K = A Q_{K-1}$ $(Q_0 = Z_0)$.

$Q_K R_K = Z_K$ (QR factorization)

$\lambda_K (Q_K^* A Q_K) = \{ \lambda_1^{(K)}, \dots \lambda_n^{(K)} \}$

(sequence of e-value estimates).

__/__/____

(Schur's theorem)

$$\underbrace{Q_k^* A Q_k}_{T_k} \rightarrow \text{an upper } \Delta r \text{ matrix}$$

We require 2 assumptions :

① The leading $(n+1)$ eigen values are distinct in abs. value

$|\lambda_1| > |\lambda_2| > \cdots |\lambda_n| > |\lambda_{n+1}| \geq \cdots \geq |\lambda_m|$

② ~~As~~ The matrix Z_0 has all non-zero principal minors.

This requirement ensures that the columns of Z_0 are not deficient in any of the directions $\{e_1 \ldots e_n\}$.

$$\langle A^i x_1^{(0)} \rangle \rightarrow \langle p_1 \rangle$$

$$\vdots$$

$$\langle A^i x_1^{(0)} \ldots A^i x_n^{(0)} \rangle \rightarrow \langle p_1 \ldots p_k \rangle.$$