Power iteration & inverse iteration.

Note Title

(I) Power iteration - this method finds the abs. value of the largest e-value of A & its corr. e-vector.

Algorithm: choose x_0 (initial vector)

$i=0$ to convergence

$$y_{i+1} = A x_i$$

$$x_{i+1} = y_{i+1} / \|y_{i+1}\|_2 \quad \text{(approx. e-vector)}$$

$$\lambda_{i+1} = x_{i+1}^T A \, x_{i+1} \quad \text{(approx. e-value)}$$

end.

$y_1 = Ax_0$

$x_1 = \frac{Ax_0}{\|Ax_0\|}$

$y_2 = A\left(\frac{Ax_0}{\|Ax_0\|}\right) = \frac{A^2x_0}{\|Ax_0\|}$

$x_2 = A^2x_0 / \|A^2x_0\|$

- Consider the simplest case where $A = \text{diag}(\lambda_1, \ldots, \lambda_n)$ with $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$

In this case the e-vectors are columns e_i of the identity matrix.

Note that x_i can be written as $x_i = \frac{A^i x_0}{\|A_i x_0\|}$.

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

Then $A^i x_0 = A^i \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \xi_1 \lambda_1^i \\ \vdots \\ \xi_n \lambda_n^i \end{pmatrix} = \xi_1 \lambda_1^i \begin{pmatrix} 1 \\ \frac{\xi_2}{\xi_1}\left(\frac{\lambda_2}{\lambda_1}\right)^i \\ \vdots \\ \frac{\xi_n}{\xi_1}\left(\frac{\lambda_n}{\lambda_1}\right)^i \end{pmatrix}$

Since each $|\lambda_i| \leq |\lambda_1|$, $A^i x_0 \longmapsto \pm \xi_1 \lambda_1^i e_1$

$\therefore \; x_i = \frac{A^i x_0}{\|A^i x_0\|} \longmapsto \pm e_1$ = e-vector corr. to largest e-value λ_1.

$x_i \mapsto \pm e_1 \implies x_i^T A x_i \mapsto \lambda_1$ = largest e-value.

- Now let A be a diagonalizable matrix i.e. $\exists$ invertible matrix P such that $A = P\Lambda P^{-1}$, $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$
$|\lambda_1| \geq \dots \geq |\lambda_n|$.

Let $P = [p_1 | \cdots | p_n]$, where $p_1, \dots, p_n$ are e-vectors with $\|p_i\|_2 = 1$.

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

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

So $A^i x_0 = P\Lambda^i P^{-1} P\begin{pmatrix}\xi_1 \\ \vdots \\ \xi_n\end{pmatrix} = P\Lambda^i \begin{pmatrix}\xi_1 \\ \vdots \\ \xi_n\end{pmatrix} = P\begin{pmatrix}\xi_1\lambda_1^i \\ \vdots \\ \xi_n\lambda_n^i\end{pmatrix} = P\,\xi_1\lambda_1^i \begin{pmatrix}1 \\ \frac{\xi_2}{\xi_1}\left(\frac{\lambda_2}{\lambda_1}\right)^i \\ \vdots \\ \frac{\xi_n}{\xi_1}\left(\frac{\lambda_n}{\lambda_1}\right)^i\end{pmatrix}$

$\therefore A^i x_0 \longmapsto \xi_1 \lambda_1^i P e_1 \qquad (Pe_1 = p_1)$

$\therefore x_i = \dfrac{A^i x_0}{\|A^i x_0\|} \longmapsto p_1$ (e-vector corr. to λ_1) $\quad$ (as $i \to \infty$)

& $\lambda_i = x_i^T A x_i \longmapsto p_1^T A p_1 = p_1^T \lambda_1 p_1 = \lambda_1$.

Remarks: (1) We need to choose x_0 such that its first component is non-zero. This is mostly okay if x_0 is chosen at random.

(2) The rate of convergence is completely dependent on the gap between λ_1 & λ_2 (the eigengap).

If $\lambda_1 >> \lambda_2$, then convergence is fast; on the other hand, if $\frac{\lambda_2}{\lambda_1}$ is close to 1, the convergence is very slow.

③ If A is real & the largest e-value is complex then there are 2 complex-conjugate e-values with the same abs. value; then this method does not work (same case if 2 real nos. k & $-k$ are e-values of A.)

(II) <u>Inverse iteration</u> (inverse power method)

useful when we know (or have a good approx, say σ) for an e-value λ and we want to find the corr. e-vector.

Idea: apply power method to a 'shift': $(A-\sigma I)^{-1}$

This makes the power method converge to an e-value closest to σ.

<u>Algorithm</u>: $i=0$ to convergence (x_0 : initial vector)

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

$$x_{i+1} = y_{i+1} / \|y_{i+1}\|_2 \qquad \text{(approx. e-vector)}$$

$$\lambda_{i+1} = x_{i+1}^T A x_{i+1} \qquad \text{(approx. e-value)}.$$

end.

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

$\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

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

$\therefore (A-\sigma I)^{-1}$ has the same e-vectors as A with corr. e-values $(\lambda_j - \sigma)^{-1}$.

Suppose $\lambda_k - \sigma \leq \lambda_i - \sigma \quad \forall i$ (i.e. λ_k is the e-value closest to σ)

i.e. $(\lambda_i - \sigma)^{-1} \leq (\lambda_k - \sigma)^{-1} \quad \forall i$

Then $(A-\sigma I)^{-i} x_0 = P(\Lambda - \sigma I)^{-i} P^{-1} P \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}$

$$= P \begin{pmatrix} \xi_1 (\lambda_1 - \sigma)^{-i} \\ \vdots \\ \xi_n (\lambda_n - \sigma)^{-i} \end{pmatrix}$$

$$= \xi_k (\lambda_k - \sigma)^{-i} P \begin{array}{cl} \begin{pmatrix} \frac{\xi_1}{\xi_k} \left(\frac{\lambda_k - \sigma}{\lambda_1 - \sigma}\right)^i \\ \vdots \\ 1 \\ \vdots \\ \frac{\xi_n}{\xi_k} \left(\frac{\lambda_k - \sigma}{\lambda_n - \sigma}\right)^i \end{pmatrix} & \begin{array}{l} \\ \\ k^{th} \text{ spot} \\ \\ \\ \end{array} \end{array}$$

Since $\left|\frac{\lambda_k - \sigma}{\lambda_i - \sigma}\right| < 1$, as $i \to \infty$, $\frac{(A-\sigma I)^{-i} x_0}{\|(A-\sigma I)^{-i} x_0\|} \longmapsto P e_k = p_k$

" e-vector corr. to λ_k.

$\therefore \; \lambda_i = x_i^T A x_i \longmapsto \lambda_k$ as $i \to \infty$