Let me emphasize main points of the power methods described in [1]. Of course, I will add some points of myself to make some concepts more clear.

1 Baisc power methods

After establishing $A^k x$ converges to the eigenvector of A corresponding to the largest eigenvalue, we need some stopping criterion, which is the main concern in numerical analysis because we cannot set max_eiter simply.

Since $x^{(m)}$ cannot be eigenvector identically, there is no real number a satisfying $Ax^{(m)} = ax^{(m)}$. But we can expect the number a which makes

$$\|Ax^{(m)} - ax\|$$

minimizes is a good approximation of the largest eigenvalue. Now we elaborate on how to solve this a.

Rayleigh Quotient. If one see $Av = av$ as an overdetermined equation system, technique about **least square** can be applied to get the solution a which is called Rayleigh quotient.

$$\Lambda_A(v) := \frac{v^* A v}{v^* v} \tag{1}$$

Geometrically, the solution of least square is just to find a scalar a such that $Av - av$ is orthogonal to v.

Rayleigh quotient

One can also deduce the same result (1) algebraically by solving the following optimization problem

$$\min_a \|Av - av\|_2^2, \tag{2}$$

Attention: one must view $f(a) = \|Av - av\|_2^2$ as a function $F(x, y)$ where $a = x + iy$ and utilize the mutiple calculus to solve the critical point.

Stopping Criterion. Combined with $\|x^{(m)} - v\| \approx 0$ and $|\lambda^{(m)} - \lambda| \approx 0$ where $\lambda, v \in \text{eig}(A)$, we can see that

$$\|Ax^{(m)} - \lambda^{(m)}x^{(m)}\| \to 0$$

where each components can be computed explicitly in the process of iteration. So one option for stopping criterion is to set ϵ and stop the iteration when

$$\|Ax^{(m)} - \lambda^{(m)}x^{(m)}\| < \epsilon \tag{3}$$

This is not impractical, for example, saying $(10^{-5}, e_1) \in \text{eig}(A)$ and we can see that

$$Ax^{(m)} - \lambda^{(m)}x^{(m)} = \begin{bmatrix} 10^{-5} + O(\epsilon) \\ \delta^{(m+1)} \\ \vdots \\ \delta^{(m+1)} \end{bmatrix} - \begin{bmatrix} \lambda^{(m)} + O(\epsilon) \\ \delta^{(m)} \\ \vdots \\ \delta^{(m)} \end{bmatrix}$$

satisfies the stopping criterion(3) after a few steps but we may not get good result. This reminds us that we need to consider the relative error, saying

$$\|Ax^{(m)} - \lambda^{(m)}x^{(m)}\| < \epsilon|\lambda^{(m)}| \tag{4}$$

2 Simultaneous Iteration

Suppose we want to find the first k eigenvectors of A, we can use the simultaneous iteration. We expect $V^{(m+1)} = A^m V$ where V is matrix consisting of k independent column vectors to converge to the eigenvectors of A corresponding to the largest k eigenvalues. Although we can prove that $V^{(m+1)}$ is injective but it's just theoretical! In practice, we need to consider the numerical stability of the algorithm. Let me give an example to show the problem

Consider the symmetric matrix:

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

We will start with the matrix V with two columns in the initial step:

$$V^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Iteration Without Orthogonalization

Each column of $V^{(m)}$ is updated via:

$$V^{(m+1)} = AV^{(m)}$$

Perform a few iterations to illustrate the convergence:

First Iteration:

$$V^{(1)} = AV^{(0)} = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}$$

Second Iteration:

$$V^{(2)} = AV^{(1)} = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 18 & 8 \\ 8 & 11 \\ 7 & 6 \end{bmatrix}$$

... 10th Iteration:

$$V^{(10)} = AV^{(9)} = \begin{bmatrix} 8490718 & 5843148 \\ 5843148 & 4032371 \\ 4458347 & 3073546 \end{bmatrix}$$

Observation

As iterations continue, each column increasingly aligns with the dominant eigenvector $\mathbf{v}_1 = \begin{bmatrix} 0.7558 \\ 0.5207 \\ 0.3971 \end{bmatrix}$ corresponding to $\lambda_{\max} = 5.2143$. For large m, the columns of $V^{(m)}$ converge to:

$$V^{(m)} \approx \begin{bmatrix} 0.756 & 0.756 \\ 0.520 & 0.520 \\ 0.397 & 0.397 \end{bmatrix}$$

i.e., all columns of $V^{(m)}$ converge to the dominant eigenvector $\mathbf{v}_1$. This convergence indicates numerical instability, as the columns lose linear independence and all vectors converge toward the dominant eigenvector's direction.

Apply QR into Simultaneous Iteration

Orthogonalize (using QR Factorization):   Perform QR factorization
when computing $V^{(2)} = AV^{(1)} = A$ where

$$V^{(1)} = \begin{bmatrix} 4 & 1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = Q^{(1)}R^{(1)} = \begin{bmatrix} -0.9428 & 0.2851 \\ -0.2357 & -0.9366 \\ -0.2357 & -0.2036 \end{bmatrix} \begin{bmatrix} -4.2426 & -1.8856 \\ 0 & -2.7285 \end{bmatrix},$$

then we compute the QR of $AQ^{(1)} = Q^2 \hat{R}^{(2)}$,

$$V^{(2)} = Q^{(2)} \hat{R}^{(2)} R^{(1)}$$

$$= \begin{bmatrix} -0.8611 & 0.4685 \\ -0.3827 & -0.8531 \\ -0.3349 & -0.2297 \end{bmatrix} \begin{bmatrix} 4.9272 & 1.3987 \\ 0 & 2.5708 \end{bmatrix} \begin{bmatrix} -4.2426 & -1.8856 \\ 0 & -2.7285 \end{bmatrix}$$

Update $V^{(m)}$ with orthogonal columns: After orthogonalization:

$$V^{(m)} = Q^{(m)} R^{(m)} = \begin{pmatrix} -0.7559 & 0.6316 \\ -0.5205 & -0.7395 \\ -0.3971 & -0.2330 \end{pmatrix} 1.0e + 07 \begin{pmatrix} -5.8556 & -4.0339 \\ 0 & -0.0024 \end{pmatrix}$$

with $m = 10$ as example. One can verify that the columns of $Q^{(10)}$ closely approximate the eigenvectors of A corresponding to its two largest eigenvalues
(in magnitude).

This orthogonalization corrects the numerical errors that lead to loss of
linear independence, thus ensuring each iteration step mainly jumps to the
invariant subspace.

Krylov space

Suppose we use Krylov space spanned by $(x, Ax, \cdots, A^{m-1}x)$ as the following
$V^{(0)}$, one can expect that the largest eigenvalue of $H_m = Q_m^* A Q_m$ approximates largest eigenvalue of A (in modulus) where Q_m is the orthogonal matrix
whose columns are the basis of $V^{(0)}$ and H_m is the Hessenberg matrix.

References

[1] Steffen Börm and Christian Mehl. *Numerical methods for eigenvalue problems*. Walter de Gruyter, 2012.