Convergence of the Power Method: Complete Proof

Setup and Assumptions

Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. We make the following assumptions:

Assumption 1 (Diagonalizability)**.** *A has n linearly independent eigenvectors $\{v_1, v_2, \dots, v_n\}$ forming a basis of $\mathbb{R}^n$.*

Assumption 2 (Distinct Dominant Eigenvalue)**.** *The eigenvalues of A can be ordered such that*

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n| \tag{1}$$

where λ_1 is called the dominant eigenvalue.

Assumption 3 (Non-zero Projection)**.** *The initial vector $v^{(0)}$ has a non-zero component in the direction of the dominant eigenvector v_1, i.e., when written as*

$$v^{(0)} = \sum_{i=1}^{n} \alpha_i v_i \tag{2}$$

we have $\alpha_1 \neq 0$.

Main Convergence Theorem

Theorem 1 (Convergence of Power Method)**.** *Under Assumptions 1-3, the normalized iterates*

$$x^{(k)} = \frac{A^k v^{(0)}}{\|A^k v^{(0)}\|} \tag{3}$$

satisfy

$$x^{(k)} \to \pm \frac{v_1}{\|v_1\|} \tag{4}$$

as $k \to \infty$, where the sign depends on the phase of λ_1^k. Moreover, the convergence is geometric with rate

$$\left\| x^{(k)} - \pm \frac{v_1}{\|v_1\|} \right\| = O\left(\left| \frac{\lambda_2}{\lambda_1} \right|^k \right) \tag{5}$$

Proof

Step 1: Eigenvalue Decomposition

Since $\{v_1, \dots, v_n\}$ form a basis (Assumption 1), we can write

$$v^{(0)} = \sum_{i=1}^{n} \alpha_i v_i \tag{6}$$

where $\alpha_1 \neq 0$ by Assumption 3.

Step 2: Applying Powers of A

By linearity and the eigenvalue equation $Av_i = \lambda_i v_i$:

$$A^k v^{(0)} = A^k \left(\sum_{i=1}^{n} \alpha_i v_i \right) \tag{7}$$

$$= \sum_{i=1}^{n} \alpha_i A^k v_i \tag{8}$$

$$= \sum_{i=1}^{n} \alpha_i \lambda_i^k v_i \tag{9}$$

Step 3: Factoring Out the Dominant Term
Factor out λ_1^k:

$$A^k v^{(0)} = \lambda_1^k \left(\alpha_1 v_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right) \tag{10}$$

Step 4: Convergence of the Direction
By Assumption 2, $\left| \frac{\lambda_i}{\lambda_1} \right| < 1$ for all $i \geq 2$. Therefore:

$$\lim_{k \to \infty} \left(\frac{\lambda_i}{\lambda_1} \right)^k = 0 \quad \text{for } i \geq 2 \tag{11}$$

This gives:

$$\lim_{k \to \infty} \frac{A^k v^{(0)}}{\lambda_1^k} = \alpha_1 v_1 \tag{12}$$

Step 5: Normalized Iterates
The normalized iterate is:

$$x^{(k)} = \frac{A^k v^{(0)}}{\|A^k v^{(0)}\|} \tag{13}$$

$$= \frac{\lambda_1^k \left(\alpha_1 v_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right)}{\left\| \lambda_1^k \left(\alpha_1 v_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right) \right\|} \tag{14}$$

$$= \frac{\lambda_1^k}{|\lambda_1|^k} \cdot \frac{\alpha_1 v_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i}{\left\| \alpha_1 v_1 + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right\|} \tag{15}$$

Let $\theta_k = \arg(\lambda_1^k)$. Then $\frac{\lambda_1^k}{|\lambda_1|^k} = e^{i\theta_k}$ (or ± 1 if λ_1 is real).
As $k \to \infty$:

$$x^{(k)} \to e^{i\theta_k} \frac{\alpha_1 v_1}{\|\alpha_1 v_1\|} = e^{i\theta_k} \frac{v_1}{\|v_1\|} \cdot \frac{\alpha_1}{|\alpha_1|} \tag{16}$$

For real matrices with real dominant eigenvalue: $x^{(k)} \to \pm \frac{v_1}{\|v_1\|}$.

Step 6: Convergence Rate Analysis
To quantify the convergence rate, consider the error:

$$A^k v^{(0)} = \lambda_1^k \alpha_1 v_1 \left(1 + \sum_{i=2}^{n} \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1}\right)^k \frac{v_i}{v_1}\right) \tag{17}$$

The dominant error term comes from λ_2:

$$\left\| x^{(k)} - \frac{v_1}{\|v_1\|} \right\| \sim \left|\frac{\alpha_2}{\alpha_1}\right| \left|\frac{\lambda_2}{\lambda_1}\right|^k \left\| \frac{v_2}{\|v_1\|} \right\| \tag{18}$$

Therefore, the convergence is geometric with ratio $\rho = \left|\frac{\lambda_2}{\lambda_1}\right|$.

Key Observations

1. **Convergence Rate:** The ratio $\rho = |\lambda_2/\lambda_1|$ determines convergence speed.
   - If $|\lambda_2| \ll |\lambda_1|$: fast convergence
   - If $|\lambda_2| \approx |\lambda_1|$: slow convergence
2. **Number of Iterations:** To reduce error by factor ϵ:

$$k \approx \frac{\log \epsilon}{\log |\lambda_2/\lambda_1|} \tag{19}$$

3. **Rayleigh Quotient:** The eigenvalue estimate

$$\lambda^{(k)} = \frac{(x^{(k)})^T A x^{(k)}}{(x^{(k)})^T x^{(k)}} \tag{20}$$

   converges to λ_1 at rate $O(|\lambda_2/\lambda_1|^{2k})$ (quadratic convergence for eigenvalue).