The Power Method

Suppose A is an $n \times n$ matrix. The *eigenvalues* of A are scalars λ which satisfy

$$A\mathbf{v} = \lambda\mathbf{v} \tag{1}$$

and the vectors $\mathbf{v}$ are the associated *eigenvectors*. Finding the eigenvalues and eigenvectors of a matrix is an important problem in many applications. In some case only the largest eigenvalue of a matrix, called the *dominant eigenvalue*, is needed. One way to find it and its associated eigenvector is with the *power method*.

The Power Method Algorithm

Given an $n \times n$ diagonalizable matrix A, a tolerance $\epsilon > 0$, and the maximum allowed number of iterations $M > 0$, the general power method algorithm can be formulated as follows.

$$
\begin{array}{ll}
\mathbf{x} := (1, 1, 1, \cdots, 1)^T & \textit{initialize eigenvector estimate} \\
\mathbf{x} := \mathbf{x}/||\mathbf{x}|| & \textit{normalize eigenvector estimate} \\
\lambda := 0 & \textit{initialize eigenvalue estimate} \\
\lambda_0 := \lambda + 2\epsilon & \textit{make sure } |\lambda - \lambda_0| > \epsilon \\
k := 0 & \textit{initialize loop counter} \\
\text{while } |\lambda - \lambda_0| \geq \epsilon \text{ and } k \leq M \text{ do} & \\
\qquad k := k + 1 & \textit{update counter} \\
\qquad \mathbf{y} := A\mathbf{x} & \textit{compute next eigenvector estimate} \\
\qquad \lambda_0 := \lambda & \textit{save previous eigenvalue estimate} \\
\qquad \lambda := \mathbf{x}^T\mathbf{y} & \textit{compute new estimate: } \lambda \approx \mathbf{x}^T A\mathbf{x} \\
\qquad \mathbf{x} := \mathbf{y} & \textit{update eigenvector estimate} \\
\qquad \mathbf{x} := \mathbf{x}/||\mathbf{x}|| & \textit{normalize eigenvector estimate} \\
\text{end while} &
\end{array}
$$

If the while-loop terminates with $k \leq M$, then we conclude the algorithm has terminated successfully. In this case λ is the dominant eigenvalue and $\mathbf{x}$ is the corresponding normalized eigenvector.

The power method's rate of convergence depends on the difference between the magnitude of the dominant eigenvalue and the other eigenvalues. Also, the power method will fail if the matrix does not have any real eigenvalues.

Derivation of the Power Method

Consider the $n \times n$ diagonalizable matrix A with linearly independent eigenvectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\ldots$, $\mathbf{v}_n$ and associated eigenvalues λ_1, λ_2, $\ldots$, λ_n, where $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$ so λ_1 is the dominant eigenvalue. Let $\mathbf{x}_0$ be given by $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + \cdots + c_n\mathbf{v}_n$ with $c_1 \neq 0$ to ensure that $\mathbf{x}_0$ has some component parallel to $\mathbf{v}_1$. Then

$$\begin{aligned} A\mathbf{x}_0 &= A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + \cdots + c_n\mathbf{v}_n) \\ &= c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + c_3A\mathbf{v}_3 + \cdots + c_nA\mathbf{v}_n \\ &= c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + c_3\lambda_3\mathbf{v}_3 + \cdots + c_n\lambda_n\mathbf{v}_n. \end{aligned}$$

Consider the iteration defined by $\mathbf{x}_{k+1} = A\mathbf{x}_k$, which can also be written as $\mathbf{x}_k = A^k\mathbf{x}_0$:

$$\begin{aligned} \mathbf{x}_k = A^k\mathbf{x}_0 &= c_1A^k\mathbf{v}_1 + c_2A^k\mathbf{v}_2 + c_3A^k\mathbf{v}_3 + \cdots + c_nA^k\mathbf{v}_n \\ &= c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + c_3\lambda_3^k\mathbf{v}_3 + \cdots + c_n\lambda_n^k\mathbf{v}_n \\ &= c_1\lambda_1^k\left[\mathbf{v}_1. + \frac{c_2}{c_1}\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{v}_2 + \frac{c_3}{c_1}\left(\frac{\lambda_3}{\lambda_1}\right)^k\mathbf{v}_3 + \cdots + \frac{c_n}{c_1}\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{v}_n\right] \end{aligned}$$

Notice that $\lim_{k\to\infty}(\lambda_i/\lambda_1)^k = 0$ for $i = 2, \ldots, n$ since λ_1 is larger in magnitude than all the other eigenvalues. All terms except $\mathbf{v}_1$ in the bracketed expression shrink to zero as $k \to \infty$ so the iteration $\mathbf{x}_{k+1} \leftarrow A\mathbf{x}_k$ converges to an eigenvector associated with the dominant eigenvalue of A. The estimated value of λ_1 can be computed with

$$\lambda_1 = \lim_{k\to\infty} \frac{\mathbf{x}_k^T A\mathbf{x}_k}{\mathbf{x}_k^T\mathbf{x}_k} = \lim_{k\to\infty} \frac{\mathbf{x}_k^T\mathbf{x}_{k+1}}{\mathbf{x}_k^T\mathbf{x}_k}.$$

If the vectors $\mathbf{x}_k$ are normalized (i.e. so $\mathbf{x}_k^T\mathbf{x}_k = 1$) as they are produced, then the dominant eigenvalue is obtained from

$$\lambda_1 = \lim_{k\to\infty} \mathbf{x}_k^T\mathbf{x}_{k+1}.$$