How to impress your friends with more power

The Power Method

A. J. Odetola[1] D. K. Reed[2]

[1]Data Science and Mathematics
Penn State University

[2]Computer Science and Mathematics
Penn State University

Math 451, Spring 2020

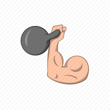


Table of Contents

Introduction to the Power Method

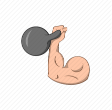
A procedure called the **power method** can be employed to compute eigenvalues. It is an example of an iterative process that, under the right conditions, will produce a sequence converging to an eigenvector of a given matrix.

Pregame

Suppose that A is an $n \times n$ matrix, and that its eigenvalues (which are unknown) have the following property:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \ldots \geq |\lambda_n|$$

We are simply ordering the eigenvalues according to decreasing absolute value. We then promote λ_1 to the **dominant eigenvalue** ;)



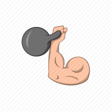
Deriving the power method

Since we know each eigenvalue has a nonzero eigenvector $\mathbf{v}_i$ from

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i \qquad (i = 1, 2, \ldots, n)$$

We assume that there is a linearly independent set of n eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. We select an arbitrary starting vector, $\mathbf{x}_0 \in \mathbb{C}^n$ and express it as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$:

$$\mathbf{x}_0 = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$$



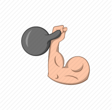
Deriving the power method

Since we know each eigenvalue has a nonzero eigenvector $\mathbf{v}_i$ from

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i \qquad (i = 1, 2, \ldots, n)$$

We assume that there is a linearly independent set of n eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. We select an arbitrary starting vector, $\mathbf{x}_0 \in \mathbb{C}^n$ and express it as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$:

$$\mathbf{x}_0 = c_1 v_1 + c_2 v_2 + \ldots + c_n v_n$$
$$A\mathbf{x}_0 = c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \ldots + c_n \lambda_n v_n$$



Deriving the power method

Since we know each eigenvalue has a nonzero eigenvector $\mathbf{v}_i$ from

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i \qquad (i = 1, 2, \ldots, n)$$

We assume that there is a linearly independent set of n eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. We select an arbitrary starting vector, $\mathbf{x}_0 \in \mathbb{C}^n$ and express it as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$:

$$\begin{aligned} \mathbf{x}_0 &= c_1 v_1 + c_2 v_2 + \ldots + c_n v_n \\ A\mathbf{x}_0 &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \ldots + c_n \lambda_n v_n \\ A^2\mathbf{x}_0 &= c_1 \lambda_1^2 v_1 + c_2 \lambda_2^2 v_2 + \ldots + c_n \lambda_n^2 v_n \end{aligned}$$

Deriving the power method

Since we know each eigenvalue has a nonzero eigenvector $\mathbf{v}_i$ from

$$A\mathbf{v}_i = \lambda_i \mathbf{v}_i \qquad (i = 1, 2, \ldots, n)$$

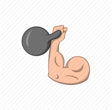
We assume that there is a linearly independent set of n eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. We select an arbitrary starting vector, $\mathbf{x}_0 \in \mathbb{C}^n$ and express it as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$:

$$\begin{aligned}
\mathbf{x}_0 &= c_1 v_1 + c_2 v_2 + \ldots + c_n v_n \\
A\mathbf{x}_0 &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \ldots + c_n \lambda_n v_n \\
A^2\mathbf{x}_0 &= c_1 \lambda_1^2 v_1 + c_2 \lambda_2^2 v_2 + \ldots + c_n \lambda_n^2 v_n \\
A^k\mathbf{x}_0 &= c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \ldots + c_n \lambda_n^k v_n
\end{aligned}$$

Deriving the power method continued

Knowing that λ_1 corresponds to the eigenvalue of greatest magnitude, we factor the **dominant eigenvalue** ;) out to get:

$$A^k \mathbf{x}_0 = \lambda_1^k \left(c_1 v_1 + c_2 v_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k + \ldots + c_n v_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \right)$$

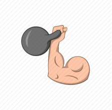


Deriving the power method continued

Knowing that λ_1 corresponds to the eigenvalue of greatest magnitude, we factor the **dominant eigenvalue** ;) out to get:

$$A^k \mathbf{x}_0 = \lambda_1^k \left(c_1 v_1 + c_2 v_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k + \ldots + c_n v_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \right)$$

As k grows towards infinity all the terms to the right of $c_1 v_1$ go to zero. since the ratios formed by factoring λ_1^k are always less than one. Thus, we see that for large values of k, our arbitrary vector x approaches an eigenvector of matrix A.



Deriving the power method continued

Knowing that λ_1 corresponds to the eigenvalue of greatest magnitude, we factor the **dominant eigenvalue** ;) out to get:

$$A^k \mathbf{x}_0 = \lambda_1^k \left(c_1 v_1 + c_2 v_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k + \ldots + c_n v_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \right)$$

As k grows towards infinity all the terms to the right of $c_1 v_1$ go to zero. since the ratios formed by factoring λ_1^k are always less than one. Thus, we see that for large values of k, our arbitrary vector x approaches an eigenvector of matrix A.

$$A^k \mathbf{x}_0 \to \lambda_1^k c_1 v_1 \qquad \text{as } k \to \infty$$

With a little more effort, we can then obtain the corresponding eigenvalue.

Table of Contents

So you couldn't handle your dominant eigenvalue and where did that bring you? Right back to λ_n

So far, we have explained how to use the power method to get the dominant eigenvector of a matrix A. But if we want to find the smallest eigenvector how would we go about doing it? Well, if we keep our previous ordering of the eigenvalues of a particular matrix A we get $|\lambda_1| \geq |\lambda_2| \geq ... > |\lambda_n| > 0$. Since A is invertible, we have that the eigenvalues for A^{-1} are,

$$|\lambda_n^{-1}| > |\lambda_n^{-1}| \geq ... \geq |\lambda_1^{-1}| > 0$$

Due to the fact that, $A\vec{X} = \lambda\vec{X} \iff A^{-1}\vec{X} = \lambda^{-1}\vec{X}$. We then proceed to perform the Power Method on A^{-1} to get λ^{-1} and by taking the reciprocal, we find the eigenvalue of least magnitude.

Table of Contents

But Devon and Ayo, what if we want the eigenvalues other than the maximum and the minimum :(

Well since you asked so nicely, lets say we want to compute an eigenvalue that is around some constant α, we apply the logic with the inverse power method by constructing an iteration where the eigenvalue we want is the dominant eigenvalue. So if we had some eigenvalues $|\lambda_j - \alpha| > \varepsilon$ and $0 < |\lambda_i - \alpha| < \varepsilon$ for $i \neq j$ and wanted to find the eigenvalue farthest away from α (i.e) a sequence that converges to $\lambda_j - \alpha$ We then have the following

$$
\begin{aligned}
A\mathbf{x} &= \lambda\mathbf{x} \\
\implies (A - \alpha I)\mathbf{x} &= (\lambda - \alpha)\mathbf{x}
\end{aligned}
$$

We then see that applying the power method to $(A - \alpha I)$ we get something that converges to $\lambda_j - \alpha$, this process is called the Shifted Power Method.

The Inverse Power Method shuffle

So now that we've found a procedure to find the the eigenvalue furthest away from some number α it only makes sense to figure out a way to get the eigenvalues closest to α (i.e) the $\lambda_i - \alpha$ this is fairly easy to accomplish from what we already know from the Shifted Power method which is

$$
\begin{aligned}
(A - \alpha I)\mathbf{v} &= (\lambda - \alpha I)\mathbf{V} \\
\implies (A - \alpha I)^{-1}\mathbf{v} &= (\lambda - \alpha)^{-1}\mathbf{v}
\end{aligned}
$$

So applying the inverse power method to $(A - \alpha I)$ gives us an approximate value for $(\lambda_i - \alpha)^{-1}$ so we just take the reciprocal and badabing-badaboom we have the eigenvalue closest to α.

Pure Art

The beautiful figure below is visualizing the Shifted Inverse Power method iterations

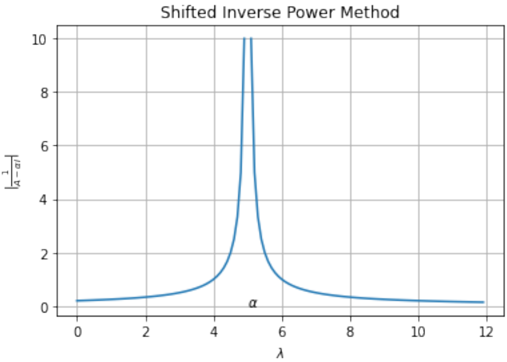
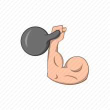


Table of Contents

Aitken Acceleration

From the sequence created by the power method, another sequence can be constructed by the means of the Aitken acceleration formula. This modified power method algorithm converges notably faster than the original sequence but for the purpose of this presentation will not covered. To learn more on Aitken Acceleration see page 363 in Kincaid and Cheney.



Questions

Does anyone have any questions?

The End

Thank You ;)