Power Method: Theory

10 points

Consider the 'shift matrix'

$$S := \begin{bmatrix} 0 & 1 & & 0 \\ & 0 & \ddots & \\ & & \ddots & 1 \\ 1 & & & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

To clarify, the following code will produce an $n \times n$ version of S:

```
import numpy as np
np.roll(np.eye(n), -1, axis=0)
```

1. Why is this called the 'shift matrix'? What does it do to a vector x when we calculate Sx?
2. Show that the eigenvalues λ of S must have absolute value $|\lambda| = 1$. That is, show that the eigenvalues are the nth roots of unity. Roots of Unity Wikipedia (https://en.wikipedia.org/wiki/Root_of_unity)
3. The power method does not converge when applied to this matrix, for any starting vector that is not already an eigenvector. Why?
4. In class we learned the conditions under which the power method converges. Why do they not apply in this case?
5. What if we applied power iteration to S^{-1}, would it converge? Please explain your answer.
6. Derive an expression for the convergence factor of power iteration on $(S - \mu I)^{-1}$, given a particular value of `k` such that S is a $k \times k$ matrix, and a particular shift, μ. Note that μ may be a complex number.

Review uploaded file (blob:https://relate.cs.illinois.edu/a6cf5923-311e-4ddd-bf4a-3fa63c29af38) · Embed viewer

Uploaded file*

选择文件 未选择任何文件

Your answer is correct.

1. It is called a shift matrix because if we apply it on a vector x, it shifts the first entry in x into the last position.

2. For any eigenvalue (λ) and eigenvector (x) pair $Sx = \lambda x$. We can take the norm of both sides of this equation in order to get $\|Sx\|_2 = \|\lambda x\|_2$. And by scalability of the norm, we get: $\|Sx\|_2 = |\lambda|\|x\|_2$. Thus: $|\lambda| = \frac{\|Sx\|_2}{\|x\|_2}$. For the shift matrix, we can say that: $\|Sx\|_2 = \|x\|_2$ since S only shifts around elements of x. And thus we conclude that: $|\lambda| = 1$.

3. The power method does not converge when applied to this matrix when the starting vector is not an eigenvector, because after applying S n times, we get back the same vector that we started with. Thus, unless we started with an eigenvector, we will keep cycling through the same vectors without making any progress.

4. In class, we said that the power method converges to the eigenvector corresponding to the largest eigenvalue in magnitude. However, there are multiple eigenvalues that are of the largest magnitude. One example is

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

which corresponds to eigenvalue 1, and another example is

$$v_2 = \begin{pmatrix} 1 \\ e^{2\pi i/n} \\ e^{(2\pi i)2/n} \\ \vdots \\ e^{2\pi i(n-1)/n} \end{pmatrix},$$

which corresponds to eigenvalue $e^{2\pi i/n}$.

5. Unshifted inverse iteration will not converge because all of the eigenvalues fall on the roots of unity, and therefore there are not any that are closer. E.g. convergence factor 1.

6. The k-th root of unity is computed as $\lambda_k = cos(\frac{2k\pi}{n}) + isin(\frac{2k\pi}{n})$. The value of k that corresponds to the numerator of the convergence factor is $p = rint(\frac{n}{2\pi} arctan(\frac{Im(\mu)}{Re(\mu)}))$, where `rint` corresponds to the round-to-the-nearest integer operator. The second closest eigenvalue to the shift is either $k+1$ or $k-1$, modulo n This can be written as $l = mod(p+1, n)$ when $atan(\frac{Im(\mu)}{Re(\mu)}) \leq \frac{p}{2\pi n}$, otherwise, $l = mod(p-1, n)$. The convergence factor is then computed as:

$$\text{\texti} convergence factor = \frac{|\mu - z_p|}{|\mu - z_l|}$$