

# Convergence of the Power Method

Michael Eden, Joshua Songy, Noam Lerner

April 19, 2016

## I Determinant, Trace, & Convergence

The power method was run against two hundred random  $2 \times 2$  matrices: the determinant and trace of each matrix was recorded. The number of iterations it took the power method to estimate the eigenvalue of the matrix within an  $\epsilon = 0.00005$  was also recorded. The results are graphed below:

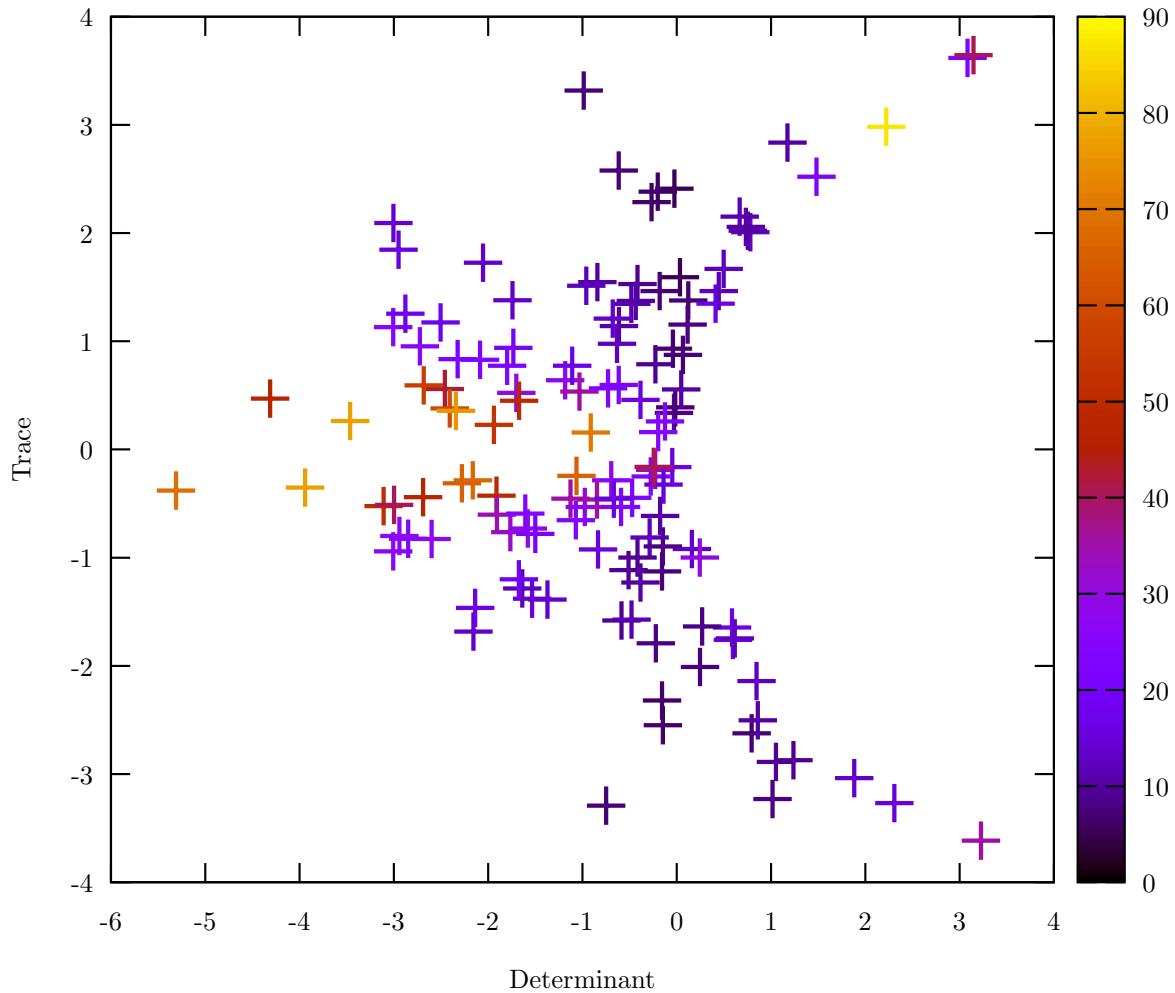


Figure 1: Determinant and Trace of random matrices vs. number of iterations of the Power Method

Clearly this graph can help us glean insight into how quickly the power method converges on a value. We know that the power method is a geometric convergence. If we let  $\lambda_1$  be the dominant eigenvalue and  $\lambda_2$  be the other one (since this is a  $2 \times 2$  matrix). Then we know the ratio of the geometric convergence is:

$$r = \left| \frac{\lambda_2}{\lambda_1} \right|$$

We can see that if  $\lambda_1$  and  $\lambda_2$  are very close together then the power method will converge more slowly (require more iterations). In the graph we see that for matrices with traces close to zero the power method took a very long time. So how does the trace and determinant relate to how close the eigenvalues are to each other? We know the two eigenvalues since they are the solution to the quadratic:

$$\begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = \lambda^2 - \lambda(a + d) + (ad - bc) = \frac{(a + d) \pm \sqrt{(a + d)^2 - 4(ad - bc)}}{2}$$

From the above note that  $(a + d)$  is a calculation for the trace of the matrix and  $(ad - bc)$  is a calculation for the determinant of the matrix. If we let  $\tau$  be the trace and  $\delta$  be the determinant then we can see a relationship between the two and the eigenvalues:

$$\frac{\tau \pm \sqrt{\tau^2 - 4\delta}}{2}$$

We can now see that we will get imaginary eigenvalues if  $4\delta > \tau^2$ . One can see a parabola with its axis of symmetry on the  $\tau$ -axis starting at the origin and growing towards the positive  $\delta$  axis that demonstrates this fact. The parabola  $4\delta = \tau^2$  marks the boundary between imaginary and real eigenvalues.

Finally, back to the problem at hand, lets see if a large  $\tau$  will affect the ratio of convergence. Lets set  $\tau = 2\alpha\sqrt{\delta}$ , so the ratio will be:

$$r = \left| \frac{\lambda_2}{\lambda_1} \right| \tag{1}$$

$$= \frac{\tau - \sqrt{\tau^2 - 4\delta}}{2} * \frac{2}{\tau + \sqrt{\tau^2 - 4\delta}} \tag{2}$$

$$= \frac{2\alpha\sqrt{\delta} - \sqrt{(2\alpha\sqrt{\delta})^2 - 4\delta}}{2\alpha\sqrt{\delta} + \sqrt{(2\alpha\sqrt{\delta})^2 - 4\delta}} \tag{3}$$

$$= \frac{2\sqrt{\delta}(\alpha - \sqrt{\alpha^2 - 1})}{2\sqrt{\delta}(\alpha + \sqrt{\alpha^2 - 1})} \tag{4}$$

$$= \frac{\alpha - \sqrt{\alpha^2 - 1}}{\alpha + \sqrt{\alpha^2 - 1}} \tag{5}$$

From the above we can clearly see that if  $\alpha = 1$  (which means  $\tau = 2\sqrt{\delta}$ ) the convergence ratio will be  $r = 1$ . However if  $\alpha$  grows to something like  $\alpha = 10$  (which means  $\tau = 20\sqrt{\delta}$ ) then  $r \approx 0.0025$ , which is **much** faster than  $\alpha = 1$ . We can now see that just like in the graph a lower trace value will result in much slower convergence times, and a higher trace value implies much faster convergence times.

A similar experiment was conducted to find the less dominant eigenvalue by running the same experiment on the inverse of each matrix in the old data. Taking the reciprocal gives the less dominant eigenvalue. We can see similar patterns in Figure 2 as we see in Figure 1, that is there is a similar boundary on the right side, and samples with smaller traces take a significantly longer time to compute below the correct  $\epsilon$ .

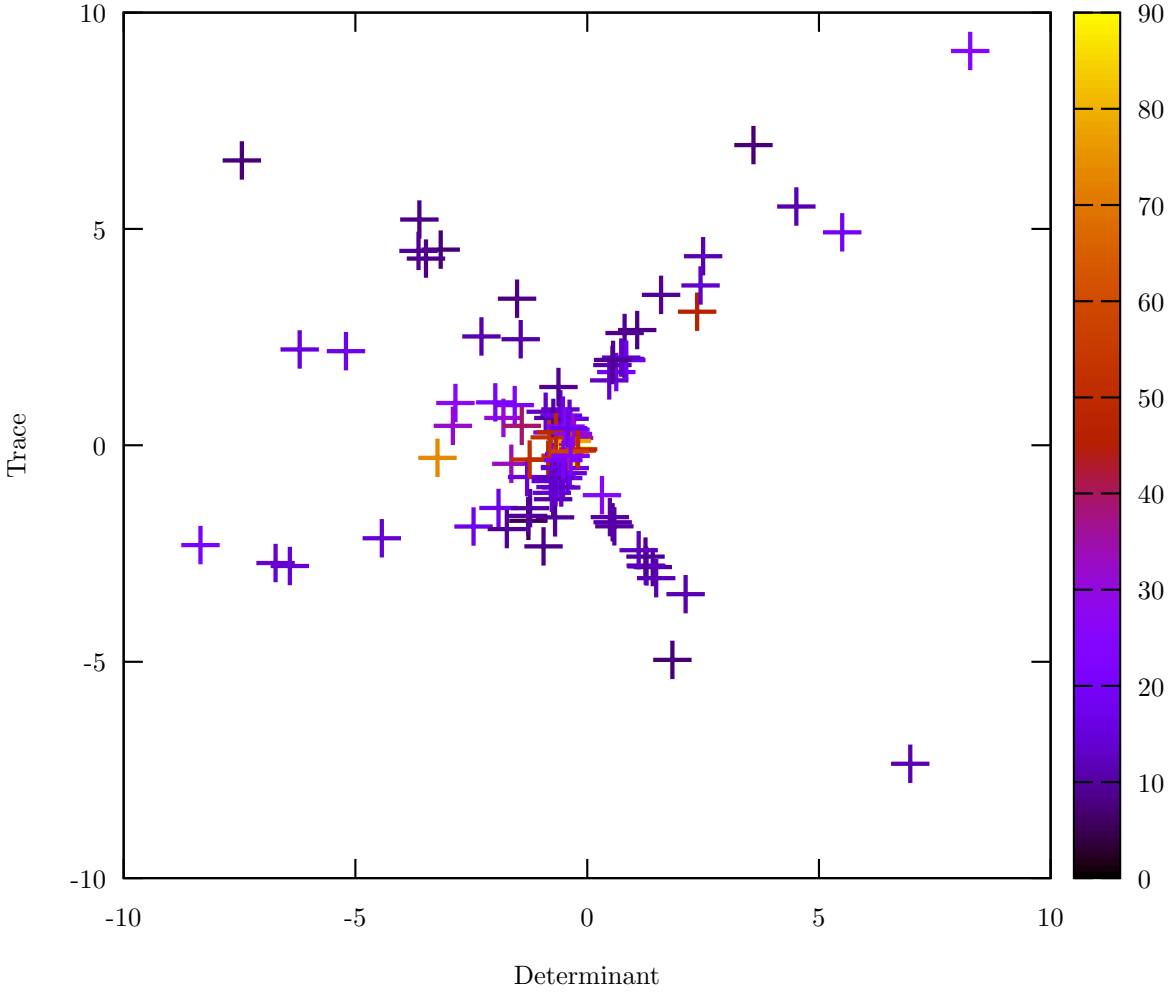


Figure 2: A similar experiment done with the inverse of the previous matrices

## II Middle Eigenvalue of a $3 \times 3$ Matrix

In this section the following matrix was examined:

$$\begin{bmatrix} -2 & 1 & 2 \\ 0 & 2 & 3 \\ 2 & 1 & -2 \end{bmatrix}$$

This matrix had eigenvalues  $\lambda_1 = -4$ ,  $\lambda_3 = -1$ . The problem is to compute the middle eigenvalue (in terms of absolute value)  $\lambda_2$ .

Two numbers were picked,  $p_1 = -(\frac{1+4}{2})$  and  $p_2 = \frac{1+4}{2}$ . Then the dominant eigenvalues for  $(A - p_1 I)^{-1}$  and  $(A - p_2 I)^{-1}$  were calculated using the power method. This really means that the inverse of the least dominant eigenvalues were calculated for  $A - p_1 I$  and  $A - p_2 I$ .

1.  $(A - p_1 I)^{-1}$  converged to  $\approx 2$
2.  $(A - p_2 I)^{-1}$  did not converge within 100 iterations.

This is a common practice to find all the eigenvalues in a matrix. When you have an algorithm that finds

one eigenvalue, you can bias the input matrix to find the others, you just have to pick the correct bias point that is far away from both eigenvalues.

This is why  $p_1$  worked so well, it was farthest away from both the smallest and largest eigenvalues (in terms of absolute value): it was the midpoint (average) of the two. This means it is also likely to be close in magnitude to the middle eigenvalue.

Since a good bias was picked we can recover the middle eigenvalue of the original matrix:

$$\lambda_{\max}((A - p_1 I)^{-1}) = 2 = \frac{1}{\lambda_2 - p_1} \implies \lambda_2 = \frac{1}{2} + p_1 \implies \lambda_2 = 3.$$

However the other bias  $p_2 = -2.5$  is shifting the matrix in the opposite way it needs to be, causing the power method to find an eigenvalue that will not be there and therefore not converge on any value.