MATH 451: Introduction to Numerical Analysis II

Instructor: Prof. Wolfgang Bangerth
bangerth@colostate.edu

Homework assignment 3 – due Wednesday 3/21/2018

Problem 1 (The power iteration). The power iteration can be used to compute the magnitude of the largest eigenvalue by magnitude, $|\lambda_1|$ of a matrix A. In class, you have seen the algorithm stated as follows: Starting with a randomly chosen initial guess $x^{(0)}$, repeat the following steps for $k = 1, 2, \ldots$:

(1) Multiply the previous iterate by A and obtain $x^{(k)} = Ax^{(k-1)}$;

(2) Compute the "amplification factor" $\mu^{(k)} = \frac{\|x^{(k)}\|_\infty}{\|x^{(k-1)}\|_\infty} = \frac{\|Ax^{(k-1)}\|_\infty}{\|x^{(k-1)}\|_\infty}$.

We have then proven that $\mu^{(k)}$ converges to $|\lambda_1|$.

The algorithm as shown works from a mathematical perspective. However, as shown in class, the iterates $x^{(k)}$ converge towards multiples of the corresponding eigenvalue – specifically, $\pm|\lambda_1|^k v_1$ (if the eigenvalues are all real, or $e^{i\varphi}|\lambda_1|^k v_1$ otherwise). The problem is that in general $|\lambda_1|^k$ either becomes very large or very small, depending on whether $|\lambda_1|$ is larger or smaller than one, and this poses problems for storing $x^{(k)}$ on computers because they can only store floating point numbers with a limited range.

Consequently, in practice one always normalizes the vector in each step. The algorithm as one would implement it then looks like as following, repeating for $k = 1, 2, \ldots$:

(1) Multiply the previous iterate by A and obtain $\tilde{x}^{(k)} = Ax^{(k-1)}$;

(2) Compute the "amplification factor" $\mu^{(k)} = \frac{\|\tilde{x}^{(k)}\|_\infty}{\|x^{(k-1)}\|_\infty} = \frac{\|Ax^{(k-1)}\|_\infty}{\|x^{(k-1)}\|_\infty}$;

(3) Set $x^{(k)} = \frac{\tilde{x}^{(k)}}{\|\tilde{x}^{(k)}\|_\infty}$.

If you think about it a bit you will realize that the two methods are the same, but that because the vector $x^{(k)}$ has unit magnitude, its elements will also be of only moderate size.

Using this *practical* algorithm, take the following sequence of matrices:

$$A_1 = (2), \qquad A_2 = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \qquad A_3 = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}, \qquad A_4 = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}, \qquad \ldots$$

These matrices are real and symmetric, and consequently we know that their eigenvalues are all in fact real. For these matrices, it can also be shown that all eigenvalues are non-negative.

(a) Use the power iteration to compute an estimate for the magnitude of the largest eigenvalue of the matrices A_{10}, A_{20}, A_{50}, and A_{100}.

(b) Let's investigate *how fast* the power iteration converges. For A_{100}, show the current value of $\mu^{(k)}$ (i.e., the current estimate of the magnitude of the largest eigenvalue) as a function of k. That is, create a graph in which you plot $\mu^{(k)}$ against k for $k = 1, \ldots, 1000$. Also state how many iterations you need until the first three significant digits of the estimate $\mu^{(k)}$ do not change any more from one iteration to the next.

(c) From parts (a) or (b), you can obtain an estimate for $|\lambda_1|$ of A_{100} if you just run the method for long enough. Since it's not the *exact* value, let's call this estimate $\tilde{\lambda}_1$. This now allows you to take the values $\mu^{(k)}$ again and plot (an approximation to the) *error* $e^{(k)} = |\mu^{(k)} - \tilde{\lambda}_1|$ as a function of k. State your estimate $\tilde{\lambda}_1$ and show such a plot. Because the errors will become quite small after some time, you will want to use a logarithmic scale for the y-axis.

(d) Verify that the relationship of $e^{(k)}$ and k indeed corresponds to linear convergence where $e^{(k)} = Ce^{(k-1)} = C^k e^{(0)}$ with a $C < 1$. What C do you infer from the data you have collected?

(50 points)

Problem 2 (Convergence acceleration). The eigenvalues of the matrices above are in fact all single (i.e., have multiplicity one). As a consequence, you have seen a theorem that states that the estimates $\mu^{(k)}$ converge linearly to $|\lambda_1|$.

Implement the convergence acceleration technique you have seen in class to compute a better estimate $\tilde{\mu}^{(k)}$ in each iteration, using $\mu^{(k-2)}$, $\mu^{(k-1)}$, and $\mu^{(k)}$. Use A_{100} from part (c) of the previous problem, and create the same kind of plot where you now show the improved guess $\tilde{\mu}^{(k)}$ against k.

For both the original data and the improved method, how many iterations do you need to obtain 3 accurate digits? **(20 points)**

Problem 3 (Multiple eigenvalues with largest magnitude). The theorem we proved in class states that the convergence of $\mu^{(k)} \to |\lambda_1|$ and $\frac{^{(k)}}{\|\tilde{x}^{(k)}\|_\infty} \to e^{i\varphi} \frac{v_1}{\|v_1\|_\infty}$ in the power iteration is only guaranteed if the largest eigenvalue by magnitude of the matrix A is genuinely larger than the next, i.e., $|\lambda_1| > |\lambda_2|$. Furthermore, we have seen that the speed of convergence is related to *how much smaller* $|\lambda_2|$ is compared to $|\lambda_1|$.

Such statements ask for an investigation of what happens if the condition is not satisfied. If you take the matrix

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

then we can write down its eigenvalues and eigenvectors by hand. This matrix, of course, has a double eigenvalue $\lambda_1 = \lambda_2 = 2$ and a smaller, single eigenvalue $\lambda_3 = 1$.

(a) Test experimentally what happens if you start the power iterations with several different random vectors. Do the $x^{(k)}$ converge to a particular direction? And does $\mu^{(k)}$ converge to a particular value?

(b) Based on the proofs you have seen in class, can you prove that we still have the convergence $\mu^{(k)} \to |\lambda_1| = |\lambda_2|$? Can you prove, based on the experience above, what the direction $x^{(k)}/\|x^{(k)}\|$ converges to?

(30 points)