Lab Session 10

MA-571 : Numerical Linear Algebra Lab 2022 R. Alam

Read very carefully.

Conditioning of eigenvalues: Let λ be an eigenvalue of $A \in \mathbb{C}^{n\times n}$ and x be an eigenvector corresponding to λ, that is, $Ax = \lambda x$ and $x \neq 0$. The eigenvector x is also called **right** eigenvector of A corresponding to λ. So, there must be something called **left** eigenvector.

A nonzero vector $y \in \mathbb{C}^n$ is called a **left** eigenvector of A corresponding to λ if $y^*A = \lambda y^*$, that is, $A^*y = \overline{\lambda} y$, where $\overline{\lambda}$ is the complex conjugate of λ. Thus for each eigenvalue λ of A there is a left eigenvector y and a right eigenvector x :

$$Ax = \lambda x \quad \text{and} \quad y^*A = \lambda y^*.$$

When λ is a simple eigenvalue of A, the quantity $\text{cond}(\lambda) := \dfrac{\|x\|_2\,\|y\|_2}{|y^*x|}$ gives the condition number of λ. If θ is the angle between x and y then it follows that $\text{cond}(\lambda) = \sec(\theta)$. Is $\text{cond}(\lambda)$ well defined? What happens when $y^*x = 0$? It is a fact that if λ is simple then y and x are unique (up to scalar multiples) and that $y^*x \neq 0$.

It is also a fact that (proof is slightly difficult) there exists a pair of left and right eigenvectors say y and x corresponding to an eigenvalue λ with $y^*x = 0$ if and only if λ is a multiple eigenvalue of A. So, $\text{cond}(\lambda) = \infty$ when λ is multiple.

Multiple eigenvalues are highly sensitive to perturbation. So, when λ is simple but $\text{cond}(\lambda)$ is large, in such a case, numerically λ behaves like a multiple eigenvalue.

Defining condition number of an eigenvector is a tricky issue because an eigenvector is NOT defined uniquely. A measure of ill-conditioning of all the eigenvectors is the quantity $\text{cond}(V) := \|V\|_2\,\|V^{-1}\|_2$, where `[V, D] = eig(A)`. We set $\text{cond}(V) = \infty$ when V is singular, that is, when A has fewer than n linearly independent eigenvectors (that is, when A is defective).

The matlab command `[V, D] = eig(A)` computes eigenvalues and eigenvectors of A. Type `help eig` for more information. You can compute condition numbers of the eigenvalues by using the command `[V, D, s] = condeig(A)`. Here `V` and `D` contain eigenvectors and eigenvalues of A, respectively, and the vector `s` contains the condition numbers of the eigenvalues, that is, `s(j)` is the condition number of the eigenvalue `D(j,j)`. Type `help condeig` for more information.

Theorem: If $\lambda(A)$ denotes a simple eigenvalue of A and E is such that $\|E\| \leq \epsilon$ for some small ϵ, then $A + E$ has a simple eigenvalue (denote that by) $\lambda(A+E)$ such that

$$|\lambda(A) - \lambda(A+E)| \leq \text{cond}(\lambda)\epsilon + \mathcal{O}(\epsilon^2).$$

*************End************

The following experiments will help you to understand sensitivity analysis of eigenvalues.

1. The MATLAB command `eig` computes eigenvalues and eigenvectors of a square matrix and the command `schur` computes Schur decomposition of a square matrix. Type help eig and schur for more information. What is the largest eigenvalue of `magic(n)` for $n = 4, 5, 6$ and why? Compute Schur decomposition of `magic(5)`.

2. The Frank matrix is a simple matrix with highly sensitive eigenvalues. The command `F = gallery('frank', n)` generates an n-by-n Frank matrix. An interesting property of Frank matrix is that if λ is an eigenvalue of F then $1/\lambda$ is also an eigenvalue of F. To see this generate a frank matrix for $n = 8$ and compute `e = eig(F)`.

   Now type `[ e'; 1 ./e']`. Can you see $(\lambda, 1/\lambda)$ pairing?

   Next compute `[ V, D, s] = condeig(A)` and type `[ diag(D)'; s']`.The last command displays eigenvalues in the first row and the corresponding condition numbers in the second row. For $n = 8, 10, 12, 14, 16$ can you see roughly how the condition numbers increase as n increases? For these values of n, what are the values of cond(V)? Do you observe some sort of relation between cond(V) and the condition numbers of the eigenvalues.

   The matrix F is Hessenberg. Now perturb F slightly as follows. Consider two perturbations $F1$ and $F2$ of F as follows. Set $F1 := F$, $F(n, 1) := 10^{-10}$ and $F2 := F$, $F2(1, n) := 10^{-10}$. For $n = 8, 10, 12$, compute eigenvalues of $F1$ and $F2$ and see if the errors satisfy the above mentioned bounds. Which perturbed matrix (among $F1$ and $F2$) shows dramatic effect of a small perturbation on the eigenvalues of F? Find cond(V) for $F1, F2$ and F. What do you observe?

   Finally, consider the frank matrix F for $n = 12$. Choose 1000 random perturbations E_i such that $\|E_i\|_2 \leq \epsilon$ and plot (in a single plot) the real and imaginary parts of the eigenvalues F and $F + E_i$ for $\epsilon := 10^{-5}, 10^{-7}, 10^{-9}$. [Type `help plot` for more info on plot command]

3. Consider the matrix given by the MATLAB command `A = gallery(5)`. Compute A^5. What are the the eigenvalues of A? Now compute eigenvalues of A using MATLAB command `eig`. What are the eigenvalues? Now plot the eigenvalues with the following commands:
```
A = gallery(5)
e = eig(A)
plot(real(e),imag(e),'r*',0,0,'ko')
axis(.1*[-1 1 -1 1])
axis square
```
   What do you observe? Next, repeat the experiment with a matrix where each element is perturbed by a single roundoff error. The elements of `gallery(5)` vary over four orders of magnitude, so the correct scaling of the perturbation is obtained with
```
e = eig(A + eps*randn(5,5).*A)
```
   Put this statement, along with the plot and axis commands, on a single line and use the up arrow to repeat the computation several times. You will see that the pentagon flips orientation and that its radius varies between 0.03 and 0.07, but that the computed eigenvalues of the perturbed problems behave pretty much like the computed eigenvalues of the original matrix.

   This experiment provides evidence for the fact that the computed eigenvalues are the exact eigenvalues of a matrix $A+E$ where the elements of E are on the order of roundoff error compared to the elements of A. This is the best we can expect to achieve with floating-point computation.

4. Now consider the Wilkinson's matrix W. It is a 20-by-20 matrix whose diagonal entries are $20, 19, \ldots, 1$, supper diagonal (just above diagonals) entries are 20 (fixed for all) and

rest of the entries are zero. This matrix can be generated as follows: `W = zeros(20); W = diag([20:-1:1])+ diag( 20 * ones(1,19), 1)`.

What are the eigenvalues of W? Compute condition number of each of the eigenvalues of W. Now perturb W slightly as follows. Set $W1 := W$ and $W1(20,1) := \epsilon$. For $\epsilon := 7.8\times 10^{-10}, 7.5\times 10^{-12}, 7.8\times 10^{-14}$, compute eigenvalues of $W1$. Do these eigenvalues satisfy the perturbation bounds $|\lambda(W) - \lambda(W1)| \leq \mathrm{cond}(\lambda)\epsilon + \mathcal{O}(\epsilon^2)$?

Now compute `[V, D] = eig(W)` and $\mathrm{cond}(V)$. Do you observe some sort of relationship between $\mathrm{cond}(V)$ and the condition numbers of the eigenvalues of W? Which eigenvalues are most sensitive to perturbations? (look at the results you have computed above)

Next, for 500 random perturbations E_i with $\|E_i\| \leq 10^{-12}$, plot (real and imaginary parts) of the eigenvalues of $W + E_i$ and W (in a single plot). The distribution of eigenvalues illustrate geometrically the sensitivity of the eigenvalues of W.

The matlab command `jordan(A)`, computes jordan canonical form of a small matrix A with integers entries. Type `help jordan` for more information. Try to compute jordan canonical forms of W and $W1$ considered above. What do you observe?

From all the results above, can you conclude that the distance of W from the set of defective matrices is $\mathcal{O}(10^{-14})$? [Exact distance is 6.13×10^{-14}.] As an illustration, compute eigenvalues of $W1$ for $\epsilon := 10^{-15}$. Then $W1$ is away from defective matrices and so W should have real and simple eigenvalues as W does. Does your experiment confirm this? Do the eigenvalues of $W1$ now satisfy the perturbation bounds given above?

5. The perturbation theory for multiple/defective eigenvalues are quite complicated. Just for illustration, consider the matrix given by the following command: `A = gallery(5)`. This is an integer matrix (that is, it has integers entries) which has many interesting properties. Matlab command `jordan(A)` computes exact jordan canonical form of A. Compute jordan canonical form of A. What are the eigenvalues of A? Now compute the eigenvalues of A by using the comand `eig(A)`. What do you observe? Next compute $\mathrm{cond}(V)$, where `[V, D] = eig(A)`. In exact arithmetic, $\mathrm{cond}(V) = \infty$. Is this fact reflected by $\mathrm{cond}(V)$? Can you explain why `eig` computes wrong eigenvalues?

   For the matrix A, the perturbation bound given above for simple eigenvalues is not valid. We will now try to guess experimentally what kind of perturbation bounds multiple eigenvalues satisfy.

   Choose random matrices E_i with $\|E_i\| = \epsilon$. For $\epsilon = 10^{-10}, 10^{-12}, 10^{-14}$, compute eigenvalues of $A + E_i$ for $i = 1 : 10$. Now compute the ratio $\frac{|\lambda(A)-\lambda_j|}{\epsilon}$, where $\lambda(A)$ is the eigenvalue of A (you know the exact eigenvalue of A from the jordan canonical form) and λ_j's are computed eigenvalue of $A + E_i$. Can you guess the value of ν for which $\frac{|\lambda(A)-\lambda_j|}{\epsilon^{1/\nu}} = \mathcal{O}(1)$ for small ϵ?

   Let J denote the jordan block of A. Set $J1 := J$ and $J1(5,1) := \epsilon$. Then $\|J - J1\| = \epsilon$. For the above values of ϵ, compute eigenvalues of $J1$ and the above ratios. For the same ν, (if you have guessed correctly), you will see that the perturbation bounds suggested by the above ratios are attained for the eigenvalues of $J1$.

************End************