Math 571 Numerical Linear Algebra Winter 2021

hw5 , due: Tuesday, March 30

1. Consider $Ax = b$. Let x be the exact solution and $\widetilde{x}$ an approximate solution. The error is $e = x - \widetilde{x}$ and the residual is $r = b - A\widetilde{x}$.

a) Show that $Ae = r$ and $\frac{||e||}{||x||} \le \kappa(A)\frac{||r||}{||b||}$.

It follows that if A is invertible, then $e = 0$ if and only if $r = 0$, but if A is ill-conditioned, then the relative error $\frac{||e||}{||x||}$ can be large even if the relative residual $\frac{||r||}{||b||}$ is small. This occurs in the following example (due to W. Kahan).

$$A = \begin{pmatrix} 1.2969 & 0.8648 \\ 0.2161 & 0.1441 \end{pmatrix},\ b = \begin{pmatrix} 0.8642 \\ 0.1440 \end{pmatrix},\ x = \begin{pmatrix} 2 \\ -2 \end{pmatrix},\ \widetilde{x}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ \widetilde{x}_2 = \begin{pmatrix} 0.9911 \\ -0.4870 \end{pmatrix}$$

b) Check that $Ax = b$ in exact arithmetic. Consider $\widetilde{x}_1, \widetilde{x}_2$ as approximate solutions and for each one compute the corresponding $\frac{||e||_\infty}{||x||_\infty}, \frac{||r||_\infty}{||b||_\infty}$, then compute $\kappa_\infty(A)$ and verify the bound in (a).

2. Consider the problem $f(x_1, x_2, x_3) = (x_1 + x_2) \times x_3$ and show that the algorithm $\widetilde{f}(x_1, x_2, x_3) = (\mathrm{fl}(x_1) \oplus \mathrm{fl}(x_2)) \otimes \mathrm{fl}(x_3)$ is backward stable. In deriving this result you may assume that each individual floating point operation is backward stable, as shown in class. (More generally, it can be shown that the composition of any two backward stable algorithms is backward stable.)

3. (a) Repeat the Matlab steps on page 41 of the lecture notes illustrating the backward stability of Householder's method for QR factorization using Matlab's `qr` command. Give the six numerical answers as in the notes. Do your results agree qualitatively with those in the notes? Discuss and explain.

b) Repeat part (a) replacing the 2nd call to Matlab's `qr` command by the modified Gram-Schmidt method using the code given below. Give the six numerical answers. Discuss and explain the similarities and differences in comparison with the results you obtained in part (a).

```
function [Q,R] = mgs(A)
[m,n] = size(A);
for i = 1:n
   v(:,i) = A(:,i);
end
for i = 1:n
   R(i,i) = norm(v(:,i),2);
   Q(:,i) = v(:,i)/R(i,i);
   for j = i+1:n
      R(i,j) = dot( Q(:,i) , v(:,j) );
      v(:,j) = v(:,j) - R(i,j) * Q(:,i);
   end
end
```

4. Solve the two-point BVP from hw1 using the compact 4th order finite-difference scheme, $-D_+D_-\left(1 - \frac{h^2}{12}d_i\right)u_i + d_i u_i = \left(1 + \frac{h^2}{12}D_+D_-\right)f_i$, with mesh size $h = 1/2^p$. Solve the linear system using the tridiagonal LU method derived in class (use vectors, not full matrices). Turn in the program listing. For $p = 1:4$ plot the exact solution ($y(x)$ vs. x) and numerical solution (u_i vs. x_i, including boundary points). For $p = 1:20$ present a table with the following data, column 1: h; column 2: $||u_h - y_h||_\infty$; column 3: $||u_h - y_h||_\infty / h^4$. Discuss and explain the trends in each column. Among the given values of h, which value ensures that the error $||u_h - y_h||_\infty$ is less than 10^{-3} for the 4th order scheme? ... for the 2nd order scheme from hw1?