Due: November 9, 2016

MATH 320: HOMEWORK 6

Please read through chapters 11, 12.1, and 13 in the textbook. Answer the following questions. Please submit all code and output with brief descriptions of what you are doing.

(1) Section 11.2 discusses various norms on matrices. In class, we reviewed three properties of any matrix norm ρ:
  (a) $\rho(aM) = a\rho(M)$ for scalar a and matrix M.
  (b) $\rho(M + N) \leq \rho(M) + \rho(N)$ for M, N matrices.
  (c) $\rho(M) = 0$ only if M is the zero matrix.

Prove that these properties are satisfied for 1) the column-sum norm, and 2) the spectral norm.
Hint: for the spectral norm, use the induced norm definition: $||M||_2 = \sup(||Mx||_2/||x||_2)$, where the norm on the right-hand side is the Euclidean norm on vectors.

Recall that the column-sum norm is given by $||M||_1 = \max_j \sum_{i=1}^{n} |M_{ij}|$. We prove each of the desired properties:
  (a) $||aM||_1 = |a|||M||_1$ for scalar a and matrix M.
  $||aM||_1 = \max_j \sum_{i=1}^{n} |aM_{ij}| = \max_j \sum_{i=1}^{n} |a||M_{ij}| = |a| \max_j \sum_{i=1}^{n} |M_{ij}| = |a|||M||_1$.
  (b) $||M + N||_1 \leq ||M||_1 + ||N||_1$ for M, N matrices.
  $||M+N||_1 = \max_j \sum_{i=1}^{n} |M_{ij}+N_{ij}|$. By the triangle inequality for real numbers, $|M_{ij} + N_{ij}| \leq |M_{ij}| + |N_{ij}|$.
  Therefore,

$$\max_j \sum_{i=1}^{n} |M_{ij} + N_{ij}| \leq \max_j \sum_{i=1}^{n} |M_{ij} + N_{ij}| \leq \max_j \left(\sum_{i=1}^{n} |M_{ij}| + \sum_{i=1}^{n} |N_{ij}| \right).$$

  If we independently maximize over each sum, it will be at least as large as maximizing them together; so:

$$\max_j \left(\sum_{i=1}^{n} |M_{ij}| + \sum_{i=1}^{n} |N_{ij}| \right) \leq \max_j \left(\sum_{i=1}^{n} |M_{ij}| \right) + \max_j \left(\sum_{i=1}^{n} |N_{ij}| \right) = ||M||_1 + ||N||_1.$$

  (c) $||M||_1 = 0$ only if M is the zero matrix.
  Suppose that M is not the zero matrix. Then some entry of the matrix is nonzero; call this entry M_{rs}. Then, $||M||_1 = \max_j \sum_{i=1}^{n} |M_{ij}|$ must be *at least* as large as $\sum_{i=1}^{n} |M_{is}| \geq |M_{rs}| \neq 0$.

We now do the same for the spectral norm:
  (a) $||aM||_2 = |a|||M||_2$ for scalar a and matrix M.
  Recall that $||M||_2 = \sup_x(||Mx||_2/||x||_2)$, so:

$$||aM||_2 = \sup_x(||aMx||_2/||x||_2) = \sup(|a|||Mx||_2/||x||_2)|a| \sup(||Mx||_2/||x||_2) = |a|||M||_2.$$

  (b) $||M + N||_2 \leq ||M||_2$ for M, N matrices.

$$||M+N||_2 = \sup_x(||(M+N)x||_2/||x||_2) = \sup_x(||Mx+Nx||_2/||x||_2) \text{ (by linearity)}$$

$$\leq \sup_x(||Mx||_2 + ||Nx||_2/||x||_2) \text{ (by triangle inequality for 2-norm)}$$

$$= \sup_x\left(\frac{||Mx||_2}{||x||_2} + \frac{||Nx||_2}{||x||_2}\right) \leq \sup_x\left(\frac{||Mx||_2}{||x||_2}\right) + \sup_x\left(\frac{||Nx||_2}{||x||_2}\right) = ||M||_2 + ||N||_2$$

(c) $||M||_2 = 0$ only if M is the zero matrix. Suppose that M is not the zero matrix. Then $M_{rs} \neq 0$ for some r, s. Take $x = e_s$, i.e. the vector which is one in the s-th coordinate and zero elsewhere. Then $Mx = M_{rs}e_r+$ an orthogonal component. For this vector $||Mx||_2/||x||_2 \geq |M_{rs}|$; therefore, the supremum over all x is at least this large. This means that $||M||_2 > 0$.

(2) Problem 11.9.

(a) Determine the condition number based on the row-sum norm for the case where $x_1 = 4, x_2 = 2, x_3 = 7$.

For this version, we need to compute explicitly by finding the row-sum norm of the matrix and its inverse.
In this case, the matrix and its inverse (computed using Gaussian elimination) are:

$$\begin{bmatrix} 16 & 4 & 1 \\ 4 & 2 & 1 \\ 49 & 7 & 1 \end{bmatrix} \qquad \begin{bmatrix} -1/6 & 1/10 & 1/15 \\ 3/2 & -11/10 & -2/5 \\ -7/3 & 14/5 & 8/15 \end{bmatrix}$$

In the first matrix, the third row is largest with sum 57; in the second matrix, the third row is largest with $\sum |A_{3j}| = 17/3 \approx 5.667$. So, the condition number based on row-sum norm is $57 \times 17/3 = 323$.

(b) Use MATLAB to compute the spectral and Frobenius condition numbers.

Here, we use the MATLAB commands written below:

```
N = [16 4 1; 4 2 1; 49 7 1];
cond(N,'fro') %Condition number based on Frobenius norm
cond(N,2)     %Condition number based on spectral norm (or induced 2-norm)
```

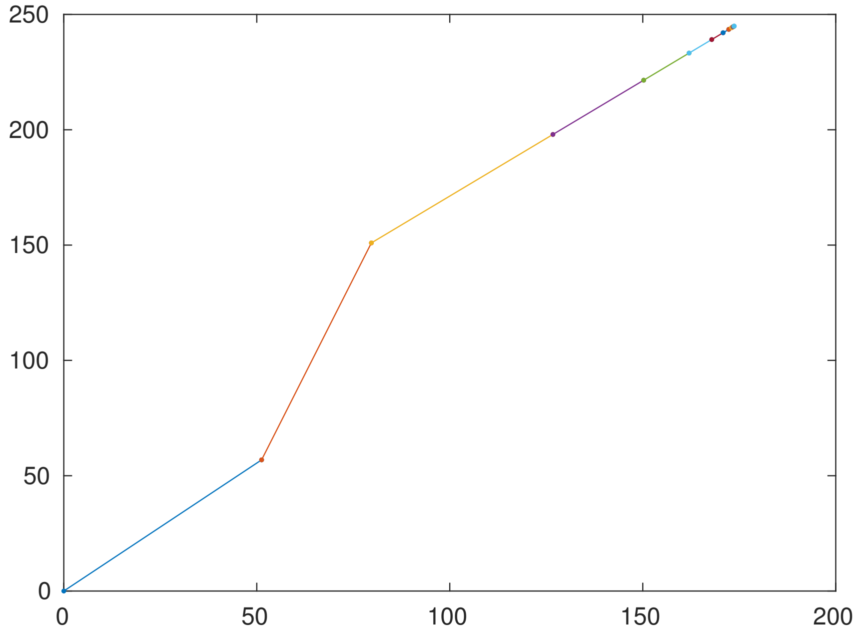
The results are: Frobenius condition number 217.4843 and spectral condition number 216.1294.

(3) Problem 12.2. You may use the textbook implementation of Gauss-Seidel, but add a subroutine that at each iteration, plots the first two coordinates of the approximation. Display the plots for part (a) and part (b).

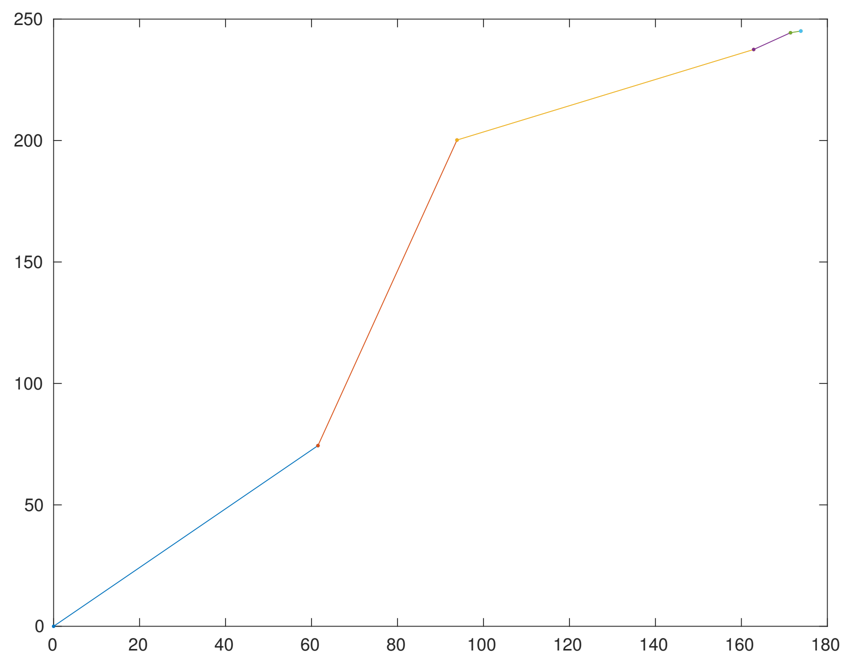
The following code is adapted from the Gauss-Seidel implementation in the textbook. The two primary changes are the inclusion of relaxation using the `lambda` parameter and the plot command in the loop.

```
function x = GaussSeidel(A,b,lambda,es,maxit)

if nargin<5|isempty(maxit),maxit=6; end
if nargin<4|isempty(es),es=0.0001;end
if nargin<3|isempty(lambda),lambda = 1;end

[m,n]=size(A);
if m ~= n, error('Matrix A must be square'); end
C = A;
x = zeros(1,n);
for i = 1:n
    C(i,i) = 0;
    x(i) = 0;
end
x = x';
for i =1:n
    C(i,1:n) = C(i,1:n)/A(i,i);
end
for i = 1:n
    d(i) = b(i)/A(i,i);
end
iter = 0;
while (iter < maxit)
    xold = x;
    for i = 1:n
        x(i) = lambda*(d(i) - C(i,:)*x) + (1-lambda)*xold(i);
        if x(i) ~= 0
            ea(i) = abs((x(i) - xold(i))/x(i)) * 100;
        end
    end
    plot([xold(1),x(1)],[xold(2),x(2)],'.-');
    hold on
    iter = iter + 1;
    if max(ea) <= es | iter >= maxit, break, end
end
disp(iter)
end
```

(a) The Gauss-Seidel method gives $[173.7041, 244.9541, 253.7270]^T$ after 13 iterations. The first two coordinates at each iteration are displayed below:



(b) The Gauss-Seidel method gives $[173.7931, 245.0580, 253.7671]^T$ after 6 iterations. The first two coordinates at each iteration are displayed below:



(4) Consider a matrix F which acts on a vector in $\mathbb{R}^2$ by mapping (x_1, x_2) to (x_2, x_1+x_2).

(a) Write F down explicitly.

The matrix sends $(1,0)$ to $(0,1)$ and $(0,1)$ to $(1,1)$; therefore, the corresponding linear transformation is:

$$F = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

(b) Let $v = [0,1]$. In a 2×11 matrix, display $F^k v$ for $k = 0, 10$.

We use the following MATLAB code to carry out this operation:

```
F = [0 1; 1 1];
v = [0 1]';

M = zeros(2,11);
for i=0:10
    M(:,i+1) = F^i*v;
end
```

The computed matrix is:

```
M =

     0     1     1     2     3     5     8    13    21    34    55
     1     1     2     3     5     8    13    21    34    55    89
```

The top row is the Fibonacci sequence, and the bottom row is the same sequence shifted by 1.

(c) What are the eigenvalues and eigenvectors of F? Use this to write an explicit formula for F^k.

To compute the eigenvalues and eigenvectors, we can use the polynomial method:

$$\begin{vmatrix} \lambda & -1 \\ -1 & \lambda - 1 \end{vmatrix} = \lambda^2 - \lambda - 1 = 0 \qquad \Rightarrow \lambda = \frac{1 \pm \sqrt{5}}{2}$$

For convenience, denote the roots by λ_+ and λ_-. The corresponding eigenvectors can be derived using the following equations:

$$\begin{pmatrix} \lambda_+ & -1 \\ -1 & \lambda_+ - 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \qquad \Rightarrow x_2 = \lambda_+ x_1$$

$$\begin{pmatrix} \lambda_- & -1 \\ -1 & \lambda_- - 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0 \qquad \Rightarrow x_2 = \lambda_- x_1$$

The corresponding eigenvectors are thus $v_+ = [1, \lambda_+]^T$ and $v_- = [1, \lambda_-]^T$. Using this formulation, the change-of-basis matrix

$$P = \begin{bmatrix} 1 & 1 \\ \lambda_+ & \lambda_- \end{bmatrix}, \qquad P^{-1} = \frac{-1}{\sqrt{5}} \begin{bmatrix} \lambda_- & -1 \\ -\lambda_+ & 1 \end{bmatrix}.$$

Using these formulae for P and P^{-1},

$$F^k = P \begin{bmatrix} \lambda_+^k & 0 \\ 0 & \lambda_-^k \end{bmatrix} P^{-1}.$$

(d) Use everything you have done so far to write a formula for the k-th Fibonacci number, where $F_0 = 0$ and $F_1 = 1$.

As we saw before, the first coordinate of $F^k v$ for $v = [0,1]^T$ is the k-th Fibonacci number. The formula for this, based on the last answer, is:

$$\begin{aligned}
&\text{first entry of } \begin{bmatrix} 1 & 1 \\ \lambda_+ & \lambda_- \end{bmatrix} \begin{bmatrix} \lambda_+^k & 0 \\ 0 & \lambda_-^k \end{bmatrix} \frac{-1}{\sqrt{5}} \begin{bmatrix} \lambda_- & -1 \\ -\lambda_+ & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \\
&= \text{first entry of } \begin{bmatrix} 1 & 1 \\ \lambda_+ & \lambda_- \end{bmatrix} \begin{bmatrix} \lambda_+^k & 0 \\ 0 & \lambda_-^k \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_+^k/\sqrt{5} \\ -\lambda_-^k/\sqrt{5} \end{bmatrix} = \frac{\lambda_+^k - \lambda_-^k}{\sqrt{5}}
\end{aligned}$$

To summarize, $F_k = \dfrac{\left(\dfrac{1+\sqrt{5}}{2}\right)^k - \left(\dfrac{1-\sqrt{5}}{2}\right)^k}{\sqrt{5}}$.

For $k \geq 2$, the power of λ_- is less than $1/2$, so it is sufficient to take the integer closest to the first term.