MATH 5610/6860: ASSIGNMENT 3

Spectral Radius: We saw that the spectral radius of a (square) matrix A never exceeds an induced matrix norm of $\|A\|$. It can be shown that for any particular matrix A one can find a vector norm such that the induced matrix norm of A is arbitrarily close to the spectral radius of A. Does the spectral radius itself define a norm? Why, or why not?

For the spectral radius to be a norm we need $r_\sigma(A) \geq 0$ for all A, and $r_\sigma(A) = 0$ if and only if A is the zero matrix. The spectral radius of any upper triangular matrix with zeros along the diagonal will be zero. Since an upper triangular matrix with zeros along its diagonal is not the zero matrix this implies the spectral radius is not a norm.

Inequalities are sharp: Explain the meaning of

$$\frac{1}{\|A\|\|A^{-1}\|}\frac{\|r\|}{\|b\|} \leq \frac{\|e\|}{\|x\|} \leq \|A\|\|A^{-1}\|\frac{\|r\|}{\|b\|}$$

and show how we derived these inequalities in class.

The inequalities above define bounds of error in an exact solution to an approximate system. Say we have a numerical solution $\hat{x}$ obtained by any numerical method to the system $Ax = b$. This solution will have some error e, i.e. $\hat{x} = x - e$. The error satisfies the same equation as the solution.

$$Ae = A(x - \hat{x}) = Ax - A\hat{x} = b - A\hat{x} = r \implies A\hat{x} = b - r$$

Where r is some residual in the system. By the properties of norms we have the following:

$$Ax = b \implies \|b\| \leq \|A\|\|x\| \tag{1}$$

$$A^{-1}b = x \implies \|x\| \leq \|A^{-1}\|\|b\| \tag{2}$$

$$Ae = r \implies \|r\| \leq \|A\|\|e\| \tag{3}$$

$$A^{-1}r = e \implies \|e\| \leq \|A^{-1}\|\|r\| \tag{4}$$

From equations (1), (2) and (4) we know

$$\frac{\|e\|}{\|A\|\|x\|} \leq \frac{\|A^{-1}\|\|r\|}{\|b\|} \implies \frac{\|e\|}{\|x\|} \leq \frac{\|A\|\|A^{-1}\|\|r\|}{\|b\|}$$

Using similar logic we can obtain the lower bound

$$\frac{\|r\|}{\|A\|\|A^{-1}\|\|b\|} \leq \frac{\|e\|}{\|x\|}$$

Combining these two results you get the inequalities above.

a. For a general matrix A, and $\|\cdot\| = \|\cdot\|_2$, show that there are non-trivial examples (i.e., $x \neq 0 \neq e$) where the right hand inequality is satisfied with equality. Do the same for the left hand inequality.

If the matrix A is orthogonal then the condition number for the 2-norm would be equal to 1. Replacing the condition number in the inequality above you end up with

$$\frac{\|r\|}{\|b\|} \leq \frac{\|e\|}{\|x\|} \leq \frac{\|r\|}{\|b\|} \implies \frac{\|e\|}{\|x\|} = \frac{\|r\|}{\|b\|}$$

b. Repeat part a. for $\|\cdot\| = \|\cdot\|_\infty$.

Similarly for the infinity norm we have equality when the condition number is equal to 1, that is equality holds whenever A has the property $\|A\|_\infty = \frac{1}{\|A^{-1}\|_\infty}$.

A symmetric matrix has real eigenvalues: Let $A \in \mathbb{R}^{n\times n}$ be *symmetric*, i.e.,

$$A = A^T.$$

Show that the eigenvalues of A are real. Hint: The appropriate generalization of a symmetric real matrix to a complex matrix is a *Hermitian Matrix* A with complex entries that satisfies $A = \bar{A}^T$ where the bar denotes conjugate complex. First show that the eigenvalues of a Hermitian matrix are real. Once you have accomplished that, try to solve the original problem without using the concept of a Hermitian Matrix. This is a great example where making a problem more general is a great help in solving it.

Proof. Letλ be some eigenvalue and x be the associated eigenvector.

$$Ax = \lambda x$$

Multiply both sides by the conjugate transpose of x

$$\bar{x}^T A x = \bar{x}^T \lambda x$$

Since λ is a scalar we can rearrange the right hand side

$$\bar{x}^T A x = \lambda(\bar{x}^T x) = \lambda\|x\|$$

Take the conjugate transpose of both sides

$$x\bar{A}^T\bar{x}^T = \bar{\lambda}\|x\|$$

Since A is Hermitian $A = \bar{A}^T$

$$x A \bar{x}^T = \bar{\lambda}\|x\|$$
$$x \lambda \bar{x}^T = \bar{\lambda}\|x\|$$
$$\lambda\|x\| = \bar{\lambda}\|x\|$$

Since $\|x\|$ is never equal to zero we can divide by it

$$\lambda = \bar{\lambda}$$

A number that is equal to it's complex conjugate must be real. Since symmetric matrices are just a special case of Hermitian matrices it follows that the eigenvalues of a symmetric matrix must also be real. $\square$

Block Matrices: Let $A \in \mathbb{R}^{m\times l}$ and $B \in \mathbb{R}^{l\times n}$. It is sometimes useful to partition matrices into blocks, e.g.,

$$A = \begin{array}{cc} & \begin{array}{cc} l_1 & l_2 \end{array} \\ \begin{array}{c} m_1 \\ m_2 \end{array} & \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \end{array} \quad \text{and} \quad B = \begin{array}{cc} & \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{array}{c} l_1 \\ l_2 \end{array} & \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \end{array}$$

where

$$l_1 + l_2 = l, \quad m_1 + m_2 = m, \qquad n_1 + n_2 = n,$$

and

$$A_{ij} \in A^{m_i \times l_j}, B_{ij} \in A^{l_i \times n_j},$$

with

$$i, j \in 1, 2.$$

The numbers outside the matrix indicate the numbers of rows and columns in part of the matrix. In this form, A and B are referred to as (2×2) *block matrices*, and the A_{ij} and B_{ij} are its *blocks*. Loosely speaking, a block matrix is a matrix whose entries are themselves matrices. Show that

$$AB = \begin{array}{cc} & \begin{array}{cc} n_1 & n_2 \end{array} \\ \begin{array}{c} m_1 \\ m_2 \end{array} & \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix} \end{array}$$

Thus, for many purposes we can treat block matrices like ordinary matrices.

Proof.

$$AB = \begin{bmatrix} a_{11} & \cdots & a_{1l} \\ a_{21} & \cdots & a_{2l} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{ml} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ b_{21} & \cdots & b_{2n} \\ \vdots & \ddots & \vdots \\ b_{l1} & \cdots & b_{ln} \end{bmatrix} = \sum_{i=1}^{l} \begin{bmatrix} a_{1i}b_{i1} & \cdots & a_{1i}b_{in} \\ a_{2i}b_{21} & \cdots & a_{2i}b_{2n} \\ \vdots & \ddots & \vdots \\ a_{mi}b_{l1} & \cdots & a_{mi}b_{ln} \end{bmatrix}$$

$$AB = \begin{bmatrix} \sum_{i=1}^{l_1} \begin{bmatrix} a_{1i}b_{i1} & \cdots & a_{1i}b_{in_1} \\ \vdots & \ddots & \vdots \\ a_{m_1 i}b_{i1} & \cdots & a_{m_1 i}b_{in_1} \end{bmatrix} & \sum_{i=l_1+1}^{l} \begin{bmatrix} a_{(l_1+1)i}b_{i(n_1+1)} & \cdots & a_{(l_1+1)i}b_{in} \\ \vdots & \ddots & \vdots \\ a_{(m_1)i}b_{i(n_1+1)} & \cdots & a_{(m_1)i}b_{in} \end{bmatrix} \\ \sum_{i=1}^{l_1} \begin{bmatrix} a_{(m_1+1)i}b_{i1} & \cdots & a_{(m_1+1)i}b_{in_1} \\ \vdots & \ddots & \vdots \\ a_{mi}b_{21} & \cdots & a_{mi}b_{2n_1} \end{bmatrix} & \sum_{i=l_1+1}^{l} \begin{bmatrix} a_{(m_1+1)i}b_{i1} & \cdots & a_{(m_1+1)i}b_{in} \\ \vdots & \ddots & \vdots \\ a_{mi}b_{21} & \cdots & a_{mi}b_{in} \end{bmatrix} \end{bmatrix}$$

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

□

More on Block Matrices, True or False:

a. Let

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

be a square matrix with square blocks A_{ij}. Is it true or false that

$$\det A = \det(A_{11})\det(A_{22}) - \det(A_{12})\det(A_{21})?$$

Proof. This is not true in general. The following is an example of where this does not work:

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

Clearly the determinant of the original matrix is 0, but if you break the matrix into blocks and calculate the determinant of the blocks it is equal to 1.

Calculating the determinant of an $n \times n$ matrix results in $n!$ terms. For example if you are dealing with a 4×4 matrix the end result would have $4! = 24$ terms. When calculating the determinant of a 4×4 using determinants of blocks you only get 8 terms. A whole bunch of terms are missing, which implies if a matrix had zeros in just the right spots this might just work.

Lets look at the 4×4 case for a general matrix A.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

If we partition this general matrix and calculate the determinant as proposed we get the following:

$$\begin{aligned} &\det(A_{11})\det(A_{22}) - \det(A_{12})\det(A_{21}) \\ &= a_{11}a_{22}a_{33}a_{44} - a_{12}a_{21}a_{33}a_{44} - a_{34}a_{43}a_{11}a_{22} + a_{12}a_{21}a_{34}a_{43} \\ &-a_{13}a_{24}a_{31}a_{42} + a_{14}a_{23}a_{31}a_{42} + a_{13}a_{24}a_{32}a_{41} - a_{14}a_{23}a_{32}a_{41} \end{aligned}$$

If we had a triangular matrix the determinant would be equal to the product of the elements along the diagonal and everything else would be zero. Calculating the determinant from its blocks would result in the same thing. □

b. Let

$$A = \begin{pmatrix} B & 0 \\ C & D \end{pmatrix}$$

where all matrices may be rectangular (i.e., not square). Is it true or false that

$$\operatorname{rank}(A) = \operatorname{rank}(B) + \operatorname{rank}(D)?$$

Proof. A counter example:

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

The rank of the original matrix is 3 and if partitioned into 2×2 matrices we have $Rank(B) + Rank(D) = 2 + 1 = 3$. It appears that it works, although look what happens when we partition it a little differently while keeping the same form with the upper right block being all zero:

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

Now we have $Rank(B) + Rank(D) = 1 + 0 = 1 \neq Rank(A)$

Working several examples it appears that only matrices that are not full rank and are partitioned in such a way that it has the maximum number of zeros in the upper right block then $Rank(B) + Rank(D) = Rank(A)$. Not sure how to prove this though. □

Operation Count: Compute how many multiplications and divisions are required to compute the Cholesky Decomposition

$$A = LL^T$$

of a symmetric and positive definite $n \times n$ matrix A, as discussed in class. (Ignore the square roots, they may be more expensive than multiplications and divisions, but there are only a few of them.)

Component	Multiplication	Division
l_{11}	0	0
l_{i1}:for $i = 2, ..., n$	(n-1)(1)	(n-1)(1)
l_{22}	(n-1)(1)	0
l_{i2}: for $i = 3, ..., n$	(n-2)(1)	(n-2)(1)
$\vdots$	$\vdots$	$\vdots$
$l_{(n-1)(n-1)}$	(n-(n-1))(1)	0
$l_{i(n-1)}$:for $i = n-1, n$	(n-(n-1))(1)	(n-(n-1))(1)
l_{nn}	1	0

This becomes

$$(n-1) + (n-2)(2) + (n-3)(3) + \cdots + (1)(n-1) = \sum_{k=1}^{n-1}(n-k)k = n\sum_{k=1}^{n-1} k - \sum_{k=1}^{n-1} k^2$$

$$\implies n\left(\frac{n(n-1)}{2}\right) - \left(\frac{n(n-1)(2n-3)}{6}\right) \approx \frac{n^3}{6}$$

The Sherman-Morrison Formula: Show that

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}$$

whenever A and $A + uv^T$ are non-singular. This is a useful formula that shows how a rank one change of a matrix amounts to a rank one change of its inverse.

Proof.

$$(A + uv^T)\left(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}\right) = A\left(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}\right) + uv^T\left(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}\right)$$

$$\implies \left(AA^{-1} - \frac{AA^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}\right) + \left(uv^TA^{-1} - \frac{uv^TA^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}\right)$$
$$\implies \left(I - \frac{uv^TA^{-1}}{1+v^TA^{-1}u}\right) + \left(uv^TA^{-1} - \frac{uv^TA^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}\right)$$
$$\implies \left(I + uv^TA^{-1} - \frac{uv^TA^{-1} + uv^TA^{-1}uv^TA^{-1}}{1+v^TA^{-1}u}\right)$$
$$\implies \left(I + uv^TA^{-1} - \frac{u(1+v^TA^{-1}u)v^TA^{-1}}{1+v^TA^{-1}u}\right)$$
$$\implies I + uv^TA^{-1} - uv^TA^{-1} = I$$

□

The Companion Matrix: Show that

$$\det\left(\begin{bmatrix} \alpha_{n-1} & \alpha_{n-2} & \cdots & \alpha_1 & \alpha_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} - \lambda I\right) = (-1)^n \left[\lambda^n - \sum_{j=0}^{n-1} \alpha_j \lambda^j\right].$$

(Thus for any polynomial p we can find a matrix whose eigenvalues are the roots of p.

Proof. Begin by expanding along first column

$$\det\begin{bmatrix} \alpha_{n-1}-\lambda & \alpha_{n-2} & \cdots & \alpha_1 & \alpha_0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \end{bmatrix}$$

$$= (\alpha_{n-1}-\lambda)\det\begin{bmatrix} -\lambda & 0 & \cdots & 0 & 0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \end{bmatrix} - (1)\det\begin{bmatrix} \alpha_{n-2} & \alpha_{n-3} & \cdots & \alpha_1 & \alpha_0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \end{bmatrix}$$

Notice that the second term is identical to the first term with the exception that it is one degree less. So we only need to see what is happening in the first term and we will know that same pattern will repeat. So, now just focusing on the first term if we expand along the first column again we have:

$$= (\alpha_{n-1}-\lambda)\left[(-\lambda)\det\begin{bmatrix} -\lambda & 0 & \cdots & 0 & 0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \end{bmatrix} - (1)\det\begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & -\lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\lambda \end{bmatrix}\right]$$

Again we notice a repeating pattern. The first term looks identical to the first term of the first level expansion except one degree less. The second term here has a row of zeros across the top which tells us that its determinant is zero so we can strike this second term. Since the pattern will continue we can replace the first level expansion with $(\alpha_{n-1}-\lambda)(-\lambda^{n-1})$. Since the second term of the first level expansion is identical to the first term we can already see we will have $n-1$ similar terms, i.e.

$$\lambda^n - \lambda^{n-1}\alpha_{n-1} + \lambda^{n-2}\alpha_{n-2} + \cdots + \lambda\alpha_1 + \alpha_0 = (-1)^n \left[\lambda^n - \sum_{j=0}^{n-1} \alpha_j \lambda^j\right]$$

$\square$

Backward Error Analysis: This problem explores the effects of a perturbation in the coefficient matrix (rather than the right hand side) of the linear system

$$Ax = b$$

Suppose we solve instead of the above system

$$(A-E)(x-e) = b$$

where E is a perturbation of A that causes an error e in the solution x. Show that

$$\frac{\|e\|}{\|x-e\|} \leq \|A\|\|A^{-1}\| \frac{\|E\|}{\|A\|.}$$

Proof.

$$(A-E)(x-e) = b$$

$$\implies \|b\| \leq \|A-E\|\|x-e\| \tag{5}$$

$$(A-E)^{-1}b = (x-e)$$

$$\implies \|x-e\| \leq \|(A-E)^{-1}\|\|b\| \tag{6}$$

$$(A-E)x - (A-E)e = b \implies (A-E)e = b - (A-E)x$$

$$\implies \|b-(A-E)x\| \leq \|A-E\|\|e\| \tag{7}$$

$$(A-E)^{-1}(b-(A-E)x) = e$$

$$\implies \|e\| \leq \|(A-E)^{-1}\|\|b-(A-E)x\| \tag{8}$$

Using 5,6, and 8 the same way we did before in question 2 of this homework:

$$\frac{\|e\|}{\|A-E\|\|x-e\|} \leq \frac{\|(A-E)^{-1}\|\|b-(A-E)x\|}{\|b\|}$$

$$\frac{\|e\|}{\|x-e\|} \leq \frac{\|A-E\|\|(A-E)^{-1}\|\|b-(A-E)x\|}{\|b\|} \leq \frac{\|A\|\|A^{-1}\|\|b-(A-E)x\|}{\|(A-E)(x-e)\|}$$

$$\implies \frac{\|e\|}{\|x-e\|} \leq \frac{\|A\|\|A^{-1}\|\|b-Ax+Ex\|}{\|Ax-Ex-Ae+Ee\|} \leq \|A\|\|A^{-1}\|\frac{\|E\|}{\|A\|}$$

$\square$