Problem 1
Use the Schur decomposition theorem to prove that, for an arbitrary norm $||\cdot||$,

$$\lim_{n\to\infty} ||A^n|| = 0 \iff \rho(A) < 1$$

Proof. $\boxed{\rightarrow}$ Suppose

$$\lim_{n\to\infty} ||A^n|| = 0$$

Note this means that A trends towards the zero matrix the larger n gets. Let v, λ^n denote an eigenvector value pair for A^n. Then

$$A^n v = \lambda^n v$$

By supposition,

$$0 = \lim_{n\to\infty} A^n v$$

and thus

$$0 = \lim_{n\to\infty} \lambda^n v$$

Since v is an eigenvector, it is nonzero. Thus it must be that

$$\lim_{n\to\infty} \lambda^n = 0$$

which can only be true if $\lambda < 1$. Notice that this holds for general eigenvalue λ, and as such we have created an upper bound. Then it must be that

$$\rho(\lambda) < 1$$

$\boxed{\leftarrow}$ Conversely, we now suppose $\rho(A) < 1$.
Using the Schur decomposition theorem, we have

$$A = QUQ^*$$

for unitary Q, and upper triangular U. Notice that this implies that

$$A^2 = QUQ^*QUQ^* = QU^2Q^*$$

and

$$A^3 = A(A^2) = QUQ^*(QU^2Q^*) = QU^3Q^*$$

Easily proved by induction, we generalize the above pattern and note

$$A^n = QU^nQ^*$$

Since U is upper triangular, any number of products with itself will result in another upper triangular matrix. That is, U^n is upper triangular for any $n \in \mathbb{N}$. Notice that U is by definition similar to A, so it has the same set of eigenvalues as A. Further, U being upper triangular means its eigenvalues must appear along the diagonal of the matrix. By supposition the the largest possible eigenvalue is still less than 1, so it must be that the diagonal entries of U are all less than 1.

$$\implies \lim_{n\to\infty} U^n = 0$$

by the nature of matrix multiplication, as all products will include a term less than 1 (meaning U^n gets arbitrarily close to zero as n gets arbitrarily large).

$$\implies \lim_{n\to\infty} QU^nQ^* = 0$$

$$\implies \lim_{n\to\infty} A^n = 0$$

Clearly any norm of the zero matrix is zero, so we may conclude

$$\lim_{n\to\infty} ||A^n|| = 0$$

as desired.

□

Problem 2
Let A be $m \times n$ and B be $n \times m$. Show that the matrices

$$\gamma = \begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \quad \text{and} \quad \beta = \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$$

have the same eigenvalues.

Proof. By definition, the eigenvalues of γ are given by

$$\det(\gamma - \lambda I) = \det\left(\begin{bmatrix} AB - \lambda I & 0 \\ B & 0 - \lambda I \end{bmatrix}\right) = 0$$

We invoke the property of block matrices which states that the determinant of a block matrix is equal to the product of the determinants of the diagonal entries. Put explicitly,

$$\det(\gamma - \lambda I) = \det\left(\begin{bmatrix} AB - \lambda I & 0 \\ B & 0 - \lambda I \end{bmatrix}\right) = \det\left(AB - \lambda I\right)\det\left(0 - \lambda I\right)$$

Notice that $\det\left(AB - \lambda I\right)$ when set equal to zero is the characteristic equation for computing the eigenvalues of AB. Similarly, we see

$$\det(\beta - \lambda I) = \det\left(\begin{bmatrix} 0 - \lambda I & 0 \\ B & BA - \lambda I \end{bmatrix}\right) = \det\left(BA - \lambda I\right)\det\left(0 - \lambda I\right)$$

Which is identical to that arrived at from γ, except we are computing the eigenvalues of BA instead of AB. Thus we've proved the claim if BA and AB have the same eigenvalues. Note these matrices are not the same dimensions, the former is an element of $\mathbb{F}^{n\times n}$ and the latter an element of $\mathbb{F}^{m\times m}$ for some field $\mathbb{F}$.

It is known that products of matrices have the same eigenvalues as the reordered product if they are in the same dimension, meaning the eigenvalues of CD and the eigenvalues of DC are the same for arbitrary matrices $C, D \in \mathbb{F}^{q\times q}$. Without loss of generality, let $m > n$. Then we'll have n possibly nonzero eigenvalues for $BA \in \mathbb{F}^{n\times n}$ (counting multiplicities). These exact n eigenvalues are shared with BA, but we have an additional $m - n$ eigenvalues of 0.

Thus the matrices γ and β share the same possibly nonzero eigenvalues, and only differ by adding additional eigenvalues of zero to which ever matrix product ordering of A and B has the higher dimension. If $m = n$, these sets will be identical, no extra zero eigenvalues needed.

$\square$

Problem 3
Show the Gershgorin theorem also holds with the bounds r_i, which are given by the partial column sums (instead of partial row sums).

Proof. Recall problem 2(c) from homework 1, which states: For $A \in \mathbb{R}^{m\times m}$, A and A^* have the same eigenvalues. That is, a real valued matrix and its transpose share the same set of eigenvalues.

Recall the Gershgorin theorem's portion of interest states the following: Let $A = \{a_{ij}\}$ be an $n \times n$ matrixd and let λ be an eigenvalue of A. Then λ belongs to one of the circles Z_i given by

$$Z_k = \{z \in \mathbb{R} \text{ or } \mathbb{C} : |z - a_{ik}| \leq r_k\}$$

where

$$r_k = \sum_{j=1, j\neq k}^{n} |a_{kj}|, k = 1, \ldots, n$$

We aim to rewrite this theorem, preserving its truth value, for r_i (the partial column sums). To that end, note

$$Av = \lambda v \implies \sum_{j=1}^{m} a_{ii} v_j = \lambda v_j$$

observe we can pull the diagonal entries out of the above summation, giving

$$\sum_{i=j}^{m} a_{ii} v_j + \sum_{i\neq j}^{n} a_{ij} v_j = \lambda v_j$$

So

$$\sum_{i=j}^{n} a_{ii} - \lambda = -\sum_{i\neq j}^{n} a_{ij}$$

which implies

$$\left| \sum_{i=j}^{n} a_{ii} - \lambda \right| = \left| -\sum_{i\neq j}^{n} a_{ij} \right|$$

and thus a single term in the left hand sum should be smaller than the whole right side (a consequence of the triangle inequality), meaning

$$|a_{ii} - \lambda| \leq \sum_{i\neq j}^{n} |a_{ij}| = r_i$$

as desired.

□

Problem 4
Use the Gershgorin theorem to show that the matrix

$$\begin{bmatrix} 1.0 & 0.3 & 0.1 & 0.4 \\ 0.0 & 2.0 & 0.0 & 0.1 \\ 0.0 & 0.4 & 3.0 & 0.0 \\ 0.1 & 0.0 & 0.0 & 4.0 \end{bmatrix}$$

has exactly one eigenvalue in each of the four circles

$$|z - k| \leq 0.1, \;\; k \in \{1, 2, 3, 4\}$$

Solution: By problem 3, we may use either the rows or the columns to bound the interval (assuming they are disjoint) that our eigenvalue falls into. I'll begin with the column form of Gershgorin's theorem, yielding that the eigenvalues must be contained in the circles:

$$z_1 : |\lambda - 1| \leq 0.1$$
$$z_2 : |\lambda - 2| \leq 0.7$$
$$z_3 : |\lambda - 3| \leq 0.1$$
$$z_4 : |\lambda - 4| \leq 0.5$$

Note all these circles are disjoint, so there should be a single eigenvalue in each. Whoa there, z_1 and z_3 immediately satisfy the given the problem statement. Let's take a look at Gershgorin's theorem in the usual row format to see if this provides any potential bounds for the remaining needed z_i. To that end,

$$z_1' : |\lambda - 1| \leq 0.8$$
$$z_2' : |\lambda - 2| \leq 0.1$$
$$z_3' : |\lambda - 3| \leq 0.4$$
$$z_4' : |\lambda - 4| \leq 0.1$$

Again, these are all disjoint and we expect to find an eigenvalue in each. Whoa doggy, z_2' and z_4' are the exact statements we needed to satisfy the goal of exactly one eigenvalue in each of the four circles

$$|z - k| \leq 0.1, \;\; k \in \{1, 2, 3, 4\}$$

Problem 5
Let $A \in \mathbb{R}^{m\times m}$ be real, symmetric and positive definite. Let $y \in \mathbb{R}^m$ be nonzero. Prove that the limit

$$\lim_{k\to\infty} \frac{y^T A^{k+1} y}{y^T A^k y}$$

exists and is an eigenvalue of A.

Proof. Recall, in general, eigenvectors that satisfy $Av = \lambda v$ also satisfy

$$A^k v = \lambda^k v$$

which follows easily by induction. Now take y to be an eigenvector. While this is a strong assumption, I believe it is necessary as the given statement does not hold for any arbitrary vector y (the limit *might* exist, but it is not necessarily an eigenvalue).

Then

$$A^{k+1} y = \lambda^{k+1} y$$
$$y^T A^{k+1} y = y^T \lambda^{k+1} y$$

Since $\lambda \in \mathbb{C}$ worst case scenario, it commutes and we may write

$$y^T A^{k+1} y = \lambda^{k+1} y^T y$$

$$\implies \lambda^{k+1} = \frac{y^T A^{k+1} y}{y^T y}$$

The exact same argument using k instead of $k+1$ gives

$$\lambda^k = \frac{y^T A^k y}{y^T y}$$

Check it out,

$$\frac{\lambda^{k+1}}{\lambda^k} = \frac{y^T A^{k+1} y}{y^T y} \left(\frac{y^T y}{y^T A^k y} \right) = \frac{y^T A^{k+1} y}{y^T A^k y}$$

where the right hand side is our original expression to be examined. Thus,

$$\lim_{k\to\infty} \frac{y^T A^{k+1} y}{y^T A^k y} = \lim_{k\to\infty} \frac{\lambda^{k+1}}{\lambda^k} = \lim_{k\to\infty} \lambda = \lambda$$

meaning the limit exists. Now we need to show λ is an eigenvalue. To that end, we search for a vector x such that

$$Ax = \lambda x$$

By construction, y will do the trick.

□

Problem 6
Let $A \in \mathbb{R}^{m\times m}$ be real with non negative entries such that

$$\sum_{j=1}^{m} A_{ij} = 1 \quad (1 \leq i \leq m).$$

Prove that no eigenvalue of A has an absolute value greater than 1.

Proof. The Girshgorin theorem will be of use here. Using the typical row set up, we are given the additional information that the whole row sums to one, and this is true for every row. Then

$$|z - a_{ii}| \leq r_i \implies -r_i \leq z - a_{ii} \leq r_i \implies -r_i + a_{ii} \leq z \leq r_i + a_{ii}$$

Since r_i is a row sum minus the diagonal entry, it follows that $r_i + a_{ii} = 1$ for all i, by the given supposition. Then

$$-r_i + a_{ii} \leq z \leq 1$$

and since this is the range of values our eigenvalues can exist in, it follows that no eigenvalue has value greater than 1. Since we're given that the entries of A are non negative, $-r_i + a_{ii}$ must be greater than -1. Then the absolute value of our eigenvalues are all no greater than 1, concluding the proof. □

Problem 7
Let $A \in \mathbb{R}^{m\times m}$ be a non-defective matrix with its eigenvalues $\{\lambda_i\}_{i=1}^m$ and its singular values $\{\sigma_i\}_{i=1}^m$ satisfying

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_m|$$

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m$$

Let $\rho(A)$ be the spectral radius of A and $\text{cond}(A) = ||A||_2||A^{-1}||_2$ be the condition number of A. Let A be normal, meaning $A^TA = AA^T$. Show that

(a) $\sigma_i = |\lambda_i|$, $1 \leq i \leq m$

(b) $||A||_2 = |\lambda_1| = \rho(A)$

(a):

Proof. We know from homework 1 problem 2(c) that the eigenvalues of a matrix and its transpose are shared. That is, if we take eigenvector/value pairing v, λ_i we have

$$Av = \lambda_i v$$

and

$$A^Tv = \lambda_i v$$

Since A is normal, $A^TA = AA^T$ and

$$A^TAv = A^T\lambda_i v = \lambda_i A^T v = \lambda_i^2 v = AA^T v$$

So the matrix $AA^T = A^TA$ has eigenvalue λ_i^2. Singular values are, by definition and the way we've constructed this set up, the square root of this eigenvalue. That is,

$$\sigma_i = \sqrt{\lambda_i^2} = |\lambda_i|$$

We have a singular value for each eigenvalue, so as desired, the above holds for $1 \leq i \leq m$.

□

(b):

Proof. By definition,

$$||A||_2 = \sqrt{\max_{\lambda}(A^*A)}$$

We are given by supposition that λ_1 is the max eigenvalue in absolute value for A, so it is also the max for A^*. Then it follows that

$$||A||_2 = \sqrt{\lambda_1^2} = |\lambda_1|$$

The spectral radius is similarly defined as the max eigenvalue in absolute value, so by definition

$$\rho(A) = |\lambda_1|$$

Putting these together, we have

$$||A||_2 = |\lambda_1| = \rho(A)$$

as desired.

□