Homework 4: Theory Questions

Julian Lehrer

Question 1. Lemma 1: $A^n = QU^nQ^*$ where U is upper triangular and Q is unitary. Proof (induction on n). For $n = 1$, we have that $A = QUQ^*$ by the Schur decomposition. Then suppose $A^n = QU^nQ^*$, and show $A^{n+1} = QU^{n+1}Q^*$. We have that $A^{n+1} = AA^n = (QUQ^*)(QU^nQ^*) = QUQ^*QU^nQ^* = QUIU^nQ^* = QU^{n+1}Q^*$, as desired. Since A^n is similar to QU^nQ^*, we have that the spectrum of A^n is the same as the spectrum of U^n. Since all norms are equivalent in a finite vector space, the result generalize for any matrix norm.

Proof ($\Longrightarrow$). Consider $\|A\|_F$, the Frobenius norm given by $\sqrt{\sum_i \sum_j |a_{ij}|}$. If $\sqrt{\sum_i \sum_j |a_{ij}^n|} \longrightarrow 0$ as $n \longrightarrow \infty$, then we must have that $|a_{ij}| < 0$, since each entry of the matrix must go to zero. Then since $\|A^n\|_F \longrightarrow 0 \iff \|U^n\|_F \longrightarrow 0$ as $n \longrightarrow \infty$ (from Lemma 1), and the diagonals of U contain the eigenvalues of A, we have by necessity that $p(A) < 1$ since $|u_{ij}| < 1$.

Proof ($\Longleftarrow$). Suppose that $p(A) < 1$. Then $p(U) < 1$, and since the diagonal elements of U are the spectra of A we have that $|u_{ii}| < 1$. Intuition: If each eigenvalue is less than one, and the determinant is the product of the eigenvalues, then $|det(A)| < 1$. Therefore, the vector mapped under A is compressed, since the determinant measures the change of volume under the image of A. Therefore, upon repeated application of a compressive map, the volume goes to zero. Since matrix norms measure how stretched a vector becomes under the image of A, it follows that the norm collapses to zero.

Proof (attempt). Since $|u_{ii}| < 1$, we have that by the Gershgorin circle theorem, all eigenvalues are within the circles centered at u_{ii} with radius $r_i = \sum_{i,i\neq j} u_{ij}$. Since we also know that the eigenvalues u_{ii} have absolute values less than 1, it follows that the radii of the Gershgorin circles must be less than $1 - u_{ii}$. But this implies that $r_i \leq 1$, so $|u_{ij}| < 1$. Therefore, the 1 norm $\max_i |\sum_{i=1}^m |u_{ij}| < 1$, so $\|U\| < 1$. Then since $\max_i |\sum_{i=1}^m |u_{ij}|^2 < \max_i |\sum_{i=1}^m |u_{ij}| < 1$, we have that $\|U^2\|_1 < \|U\|_1 < 1$. Therefore, $\|U^n\|_1 \longrightarrow 0$ as $n \longrightarrow \infty$.

Question 2. Lemma 1:

$$\det \begin{pmatrix} A & 0 \\ B & C \end{pmatrix} = \det(A)\det(C) = \det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

Proof: See https://www.statlect.com/matrix-algebra/determinant-of-block-matrix.

I wasn't fully able to prove this myself, although I learned it in undergraduate linear algebra.

Now, note that if two matrices have the same characteristic polynomial, then necessarily they have the same eigenvalues. Therefore, consider

$$\det \begin{pmatrix} AB - \lambda I & 0 \\ B & -\lambda I \end{pmatrix} = \det(AB - \lambda I)\det(-\lambda I)$$

and

$$\det \begin{pmatrix} -\lambda I & 0 \\ B & BA - \lambda I \end{pmatrix} = \det(-\lambda I)\det(BA - \lambda I)$$

Finally, we show that $\det(BA-\lambda I) = \det(AB-\lambda I)$. Equivalently, that AB and BA have the same eigenvalues. Suppose λ is an eigenvalue of AB with eigenvector x. Then $ABx = \lambda x \iff BABx = B(\lambda x)$. Then letting $y = Bx$, we have that $BAy = \lambda y$, so λ is an eigenvalue of BA. The characteristic polynomials are necessarily the same, so $\det(BA - \lambda I) = \det(AB - \lambda I)$, which completes the proof.

Question 3. First, note that since $\det A = \det A^T$, we have that $\det(A-\lambda I) = \det((A - \lambda I)^T) = \det(A^T - \lambda I)$, so the eigenvalues of A and A^T are the same. Therefore, the Gersgorin circles defined by the rows of A^T (columns of A) contain all eigenvalues of A. Equivalently, the theorem holds with column sums.

Question 4. First, consider the absolute row sums given by $r_{1,2,3,4} = 0.8, 0.1, 0.4, 0.1$. Then since we showed the Gershgorin discs can also be found by considering the absolute column sums, consider the column sums of columns 2 and 4, given by $c_{2,4} = 0.1, 0.1$. Therefore, the radius of each circle is 0.1. Additionally, since $k + 0.1 < (k + 1) - 0.1$ the circles are disjoint, and we can conclude that there is exactly one eigenvalue in $|z - k| < 0.1$ for $k = 1, 2, 3, 4$.

Question 5. Lemma: If $Ay = \lambda y$, then $A^n y = \lambda^n y$. Proof (by induction on n). $n = 1$ is handled in the definition. Then suppose $A^n y = \lambda^n y$ and show that $A^{n+1}y = \lambda^{n+1}y$. Then $A^{n+1} = AA^n = A(\lambda^n y) = \lambda^n(Ay) = \lambda^n\lambda = \lambda^{n+1}$. Then we have that $y^T A^k y = y^T y\lambda^k$, so

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

Is an eigenvalue of A.

Now note that since A is non-defective, A is diagonalizable and therefore has an eigenbasis. Then consider an arbitray vector $y = \sum_{i=1}^{m} a_i v_i$ where v_i is the ith eigenvector with corresponding eigenvalue λ_i. Then from our Lemma, we have that

$$A^n y = \sum_{i=1}^{m} \lambda_i^n a_i v_i$$

And therefore

$$\frac{y^T A^{n+1} y}{y^T A^n y} = \frac{(\sum_{i=1}^m a_i v_i^T)(\sum_{i=1}^m \lambda_i^{n+1} a_i v_i)}{(\sum_{i=1}^m a_i v_i^T)(\sum_{i=1}^m \lambda_i^n a_i v_i)} = \frac{\sum_{i=1}^m \lambda_i^{n+1} a_i v_i}{\sum_{i=1}^m \lambda_i^n a_i v_i}$$

Since A is positive definite, we know that the eigenvalues are positive ($Ax = \lambda x \implies x^T Ax = \|x\|\lambda > 0$). Let $\lambda_1 \geq ... \geq \lambda_m$. Therefore,

$$\frac{\sum_{i=1}^m \lambda_i^{n+1} a_i v_i}{\sum_{i=1}^m \lambda_i^n a_i v_i} = \frac{\lambda_1 a_1 v_1 + ... + \lambda_m \left(\frac{\lambda_m}{\lambda_1}^n\right) a_m v_m}{a_1 v_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^n a_2 v_2 + ... + \left(\frac{\lambda_m}{\lambda_1}\right)^n a_m v_m}$$

Where we divide the numerator and denominator by λ_1^n. Note that since $p(A) = \lambda_1$, $\lambda_i/\lambda_1 < 1$ for $i = 2, ..., m$. Then we have that as $n \longrightarrow \infty$ the quotient converges to λ_1.

Question 6. Lemma: $p(A) \leq \|A\|$. Proof: We consider the proof with the 2-norm given by $\|A\|_2 = \sup_{\|x\|=1} \|Ax\|_2$, since all norms are equivalent in a finite vector space. Let $p(A) = |\lambda|$, and let the corresponding eigenvector be x with $\|x\| = 1$. Then $\|Ax\| = \|\lambda x\| = |\lambda|$. Now consider an arbitary unit vector u. By the Cauchy-Schwartz inequality, we have that $\|Au\| \leq \|A\|\|u\| = \|A\|$, therefore $\|A\| \geq \|Au\|$ for all vectors u. In particular, $\|A\| \geq \|Ax\| = |\lambda|$, so $p(A) \leq \|A\|$.

Now, consider the fact that since A has nonnegative entries, $\sum_j^m a_{ij} = 1 = \|A\|_1$, the 1-norm of A. Therefore, $p(A) < 1$, or equivalently, no eigenvalue has an absolute value greater than one.

Question 7.

a. Consider the SVD of A to be $A = U\Sigma V^T$. Then since $A^T = V\Sigma U^T$, and $A^T A = V\Sigma^2 V^T = AA^T = U\Sigma^2 U^T$, we have that $U = T$. Let σ_i be the ith singular value of A. Since $\sigma_i = \sqrt{\lambda_i(A^T A)}$, $A = U\Sigma U^T$ and $A^T A = U\Sigma^2 U^T$ by virtue of A being normal, $\sigma_i = \sqrt{\sigma_i^2} = |\lambda_i|$.

b. Since $\|A\|_2 = \sqrt{p(A^T A)} = \sigma_{\max}(A)$ by definition, and we just showed that $\sigma_i = |\lambda_i|$, we have that $\|A\|_2 = |\lambda_i|$.