Math 4310 Name: Rami Pellumbi (rp534)

Homework 8

Please print out these pages. I encourage you to work with your classmates on this homework. Please list your collaborators on this cover sheet. (Your grade will not be affected.) Even if you work in a group, you should write up your solutions yourself! You should include all computational details, and proofs should be carefully written with full details. As always, please write neatly and legibly.

Please follow the instructions for the "extended glossary" on separate paper (LaTeX it if you can!) Hand in your final draft, including full explanations and write your glossary in complete, mathematically and grammatically correct sentences. Your answers will be assessed for style and accuracy.

Please **staple** *this cover sheet, your exercise solutions, and your glossary together, in that order, and hand in your homework in class.*

GRADES

Exercises ______________________ / 50

Extended Glossary

Component	Correct?	Well-written?
Definition	/6	/6
Example	/4	/4
Non-example	/4	/4
Theorem	/5	/5
Proof	/6	/6
Total	/25	/25

Exercises.

1. Suppose $T \in \mathcal{L}(V)$ is invertible, where $\dim V = n$ is a vector space over a field $\mathbb{F}$, and that the distinct eigenvalues of T are $\lambda_1, ..., \lambda_m$. Also suppose that the eigenspace E_{λ_i} has dimension d_i for each i.

   Find the eigenvalues of T^{-1}, and the dimension of its eigenspace at each eigenvalue.

   λ_i an eigenvalue of T means that
   $$T\vec{v_i} = \lambda_i \vec{v_i}.$$
   Apply T^{-1} to both sides (know that T is invertible) to get
   $$T^{-1}T\vec{v_i} = T^{-1}\lambda_i \vec{v_i}.$$
   simplify to obtain
   $$T^{-1}\vec{v_i} = \frac{1}{\lambda_i}\vec{v_i}.$$
   Thus, T^{-1} has the same amount of distinct eigenvalues as T, with the eigenvalues of T^{-1} being the eigenvalues of T inverted.

   i.e. The distinct eigenvalues of T^{-1} are $\frac{1}{\lambda_1}, ..., \frac{1}{\lambda_m}$.

   The eigenspace of T corresponding to λ_i, for some i, is defined as
   $$E_{\lambda_i} = ker(T - \lambda_i I).$$
   and it consists of the set of all eigenvectors of T corresponding to λ_i as well as the zero vector.

   We are given that
   $$dim(E_{\lambda_i}) = d_i.$$

From here it is easy to see that the eigenspace of T^{-1} corresponding to $\frac{1}{\lambda_i}$, for some i, is defined as

$$E_{\frac{1}{\lambda_i}} = \ker(T - \frac{1}{\lambda_i}I),$$

and it consists of the set of all eigenvectors of T^{-1} corresponding to $\frac{1}{\lambda_i}$ as well as the zero vector.

But the eigenvectors corresponding to each λ_i are the same as the eigenvectors corresponding to $\frac{1}{\lambda_i}$ as shown above. Thus, the dimension of E_{λ_i} is the same as the dimension of $E_{\frac{1}{\lambda_i}}$ for all i.

2. Find $A, B \in \mathbb{R}^{4\times 4}$ such that each A and B have $-1, 3, 10$ as eigenvalues, and they have no other eigenvalues, and such that A and B are not similar.

   Duplicate the eigenvalue -1 and write A as the diagonal matrix

$$A = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix}$$

   Similarly, if we instead duplicate the eigenvalue three and write B as the diagonal matrix

$$B = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 10 \end{pmatrix}$$

   We have satisfied the criteria. It is easy to see that A is not similar to B because if it was they would have the same trace.

3. Consider the stochastic matrix

$$A = \begin{pmatrix} 0.6 & 0.1 & 0.1 \\ 0.1 & 0.9 & 0.2 \\ 0.3 & 0 & 0.7 \end{pmatrix}$$

   (a) Find the Gershgorin disks for A.
   Given A, the i^{th} Gershgorin disk is

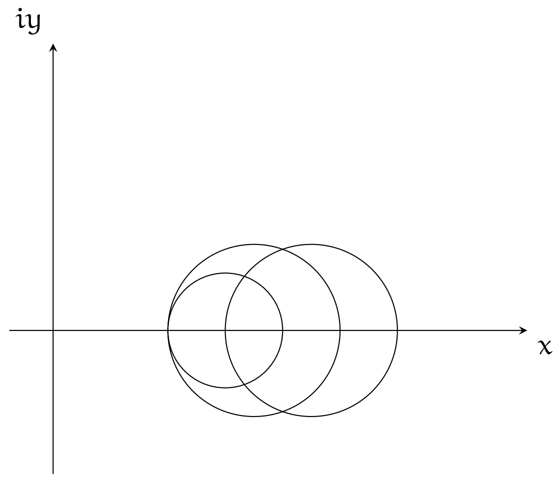
$$C_i = \{z \in \mathbb{C} \mid |z - A_{ii}| \leq r_i\}$$

   where $r_i = \rho_i(A) - |A_{ii}|$.
   We will thus have three Gershgorin disks. First we compute the sum of absolute values in the three rows to obtain $\rho_1(A) = 0.8$, $\rho_2(A) = 1.2$, $\rho_3(A) = 1$.
   Next we write the set form for the Gershgorin disks.

$$C_1 = \{z \in \mathbb{C} \mid |z - 0.6| \leq 0.2\}$$

$$C_2 = \{z \in \mathbb{C} \mid |z - 0.9| \leq 0.3\}$$
$$C_3 = \{z \in \mathbb{C} \mid |z - 0.7| \leq 0.3\}$$

So the first Gershgorin disk is a circle of radius 0.2 centered at $z = 0.6$,
the second Gershgorin disk is a circle of radius 0.3 centered at $z = 0.9$,
and the the third Gershgorin disk is a circle of radius 0.3 centered at $z = 0.7$.
This can be seen in the figure below.



(b) Find the eigenvalues for A.

Know that the eigenvalues of A will like in a Gershgorin disk, and that they will have magnitude less than 1.

The eigenvalues of A are such that

$$\lambda_1\lambda_2\lambda_3 = 0.35$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 2.2$$
$$\lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 1.74$$

Thus,

$$\lambda_1 = 0.5, \lambda_2 = 0.7, \lambda_3 = 1$$

(c) If A is the transition matrix for a 3 state Markov chain, and the initial probability vector(at time step 0) is

$$P = \begin{pmatrix} 0.3 \\ 0.3 \\ 0.4 \end{pmatrix},$$

find the proportions in each state after 2 time steps. Find the eventual proportions in each state, (i.e. the limiting proportions, as the time step goes to infinity), by computing the fixed probability vector.

After 2 time steps, the proportions are

$$A * A * P = \begin{pmatrix} 0.225 \\ 0.441 \\ 0.334 \end{pmatrix}$$

Next, we note that

$$L = \lim_{n\to\infty} A^n = \begin{pmatrix} 0.2 & 0.2 & 0.2 \\ 0.6 & 0.6 & 0.6 \\ 0.2 & 0.2 & 0.2 \end{pmatrix}$$

(Note: it makes sense that all the columns are the same since the eigenspace of A for each eigenvalue is 1)

Calculating the steady-state probability vector $L * P$, we obtain

$$L * P = \begin{pmatrix} 0.2 \\ 0.6 \\ 0.2 \end{pmatrix}$$

4. Let V be a vector space of dimension n over a field $\mathbb{F}$, and throughout this entire problem,let $S, T \in \mathcal{L}(V)$ be diagonalizable linear operators.

   (a) Show the following: if there is a basis $\mathcal{B}$ of V such that both $[S]_{\mathcal{B}}$ and $[T]_{\mathcal{B}}$ are diagonal matrices, then S and T commute: ST = TS.
   Let $[S]_{\mathcal{B}}$ be the $n \times n$ diagonal matrix

$$[S]_{\mathcal{B}} = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$$

   and $[T]_{\mathcal{B}}$ be the $n \times n$ diagonal matrix

$$[T]_{\mathcal{B}} = \begin{pmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{pmatrix}.$$

   Then

$$[S]_{\mathcal{B}}[T]_{\mathcal{B}} = \begin{pmatrix} d_1c_1 & & \\ & \ddots & \\ & & d_nc_n \end{pmatrix}$$

   and

$$[T]_{\mathcal{B}}[S]_{\mathcal{B}} = \begin{pmatrix} c_1d_1 & & \\ & \ddots & \\ & & c_nd_n \end{pmatrix}$$

   Since $c_id_i = d_ic_i$ for all i, $ST = TS$ as desired.

(b) Suppose that $U \subset V$ is T-invariant, and that $v_1, ..., v_k$ are eigenvectors with respect to different eigenvalues. Show that if $v_1 + ... + v_k \in U$, then each $v_i \in U$, for i=1..k.

$v_1, ..., v_k$ are eigenvectors of T with corresponding eigenvalues $\lambda_1, ..., \lambda_k$ and these eigenvalues are all distinct.

If

$$v_1 + ... + v_k \in U,$$

then

$$T(v_1 + ... + v_k) = T(v_1) + ... + T(v_k) \in U$$

since U is T-invariant. But these v are eigenvectors of T. i.e.

$$T(v_1) + ... + T(v_k) = \lambda_1 v_1 + ... + \lambda_k v_k \in U$$

To recap, we have

$$v_1 + ... + v_k \in U, \tag{1}$$

and the linear transformation T performed on (1)

$$\lambda_1 v_1 + ... + \lambda_k v_k \in U \tag{2}$$

Consider the following process where we multiply (1) by λ_k and subtract the result from (2).

$$(\lambda_1 - \lambda_k)v_1 + ... + (\lambda_{k-1} - \lambda_k)v_{k-1} \in U \tag{3}$$

note that this result is ALSO in U.

Now consider repeating this process in a similar fashion. Take the linear transform T of (3), then subtract (3) multiplied by λ_{k-1}. This result will also be in U as shown below.

$$\begin{aligned}
&T((\lambda_1 - \lambda_k)v_1 + ... + (\lambda_{k-1} - \lambda_k)v_{k-1}) - \lambda_{k-1}((\lambda_1 - \lambda_k)v_1 + ... + (\lambda_{k-1} - \lambda_k)v_{k-1}) \\
&= T((\lambda_1 - \lambda_k)v_1) + ... + T((\lambda_{k-1} - \lambda_k)v_{k-1}) - \lambda_{k-1}((\lambda_1 - \lambda_k)v_1 + ... + (\lambda_{k-1} - \lambda_k)v_{k-1}) \\
&= (\lambda_1 - \lambda_k)\lambda_1 v_1 + ... + (\lambda_{k-1} - \lambda_k)\lambda_{k-1}v_{k-1} - \lambda_{k-1}((\lambda_1 - \lambda_k)v_1 + ... + (\lambda_{k-1} - \lambda_k)v_{k-1}) \\
&= ((\lambda_1 - \lambda_k)\lambda_1 - \lambda_{k-1})v_1 + ... + ((\lambda_{k-2} - \lambda_k)\lambda_{k-2} - \lambda_{k-1})v_{k-2} \\
&= (\lambda_1^2 - \lambda_k\lambda_1 - \lambda_{k-1})v_1 + ... + (\lambda_{k-2}^2 - \lambda_k\lambda_{k-2} - \lambda_{k-1})v_{k-2} \in U
\end{aligned}$$

This result is ALSO in U.

Now stay with me. Consider the k^{th} step of this non intuitive process. It's difficult to imagine, I know, but the result becomes

$$(\lambda_1^k - \lambda_k\lambda_1^{k-1} - \lambda_{k-1}\lambda_1^{k-2} - ... - \lambda_2)v_1 \in U$$

.

Thus, $v_1 \in U$. We now peel the onion back so that working backwards it is easy to see $v_1 + v_2 \in U$,...,$v_1 + ... + v_k \in U$.

(c) If $U \subset V$ is T-invariant, then show that the induced map $T|_U : U \to U$ is also diagonalizable.
Need to find a basis of eigenvectors $\mathcal{A}$ for U so that $[T_{|_U}]_{\mathcal{A}}$ is diagonalizable.
If T has n distinct eigenvalues with n distinct corresponding eigenvectors, know that

$$V = E_{\lambda_1} \oplus ... \oplus E_{\lambda_n}.$$

Some subset of these eigenspaces comprises U. Note that it need not be every eigenvalue distinct, and that eigenvalues eigenspace instead needs to have dimension equal to the multiplicity.
Doing so you can take the eigenvalues corresponding to that eigenspace and realize they span your space as shown in part b. They are also linearly independent and thus they form a basis for U. Thus T restricted to U is diagonalizable as desired.

(d) Show the converse of part (a): If $ST = TS$, then there is a basis $\mathcal{B}$ of V such that both $[S]_{\mathcal{B}}$ and $[T]_{\mathcal{B}}$ are diagonal matrices, (Remember, in all parts of this problem: S and T are assumed to be diagonalizable! Also, as a hint, you might find the earlier parts of the problem could help with later parts).
This question was left as an exercise for the reader due to an exam worth a lot percent of the class grade I must take the day this is due.

5. The vector space of polynomials $\mathbb{F}[x]$ has a multiplication on it too (we can multiply 2 polynomials, and get a new polynomial. This makes $\mathbb{F}[x]$ into what is called an algebra). An ideal is a subspace $I \subseteq \mathbb{F}[x]$ which satisfies the property that if $h(x) \in \mathbb{F}[x]$, and $f(x) \in I$, then $h(x)f(x) \in I$.

(a) Fix a polynomial f(x). Show that the set

$$\langle f(x) \rangle = \{h(x)f(x) \in \mathbb{F}[x] \mid h(x) \in \mathbb{F}[x]\}$$

is an ideal in $\mathbb{F}[x]$ (called the *ideal generated by* $f(x)$).
For fixed f, it is clear that the product $h(x)f(x) \in \mathbb{F}[x]$. In fact for all possible $h(x)$ it will either comprise the entirety of $\mathbb{F}[x]$ OR some subset $I \subseteq \mathbb{F}[x]$.
Thus by definition $\langle f(x) \rangle \in I$ and is thus an ideal in $\mathbb{F}[x]$.

(b) Suppose that the ideal I contains a polynomial $g(x)$ of degree d, but no nonzero polynomial of lower degree. Show that every polynomial in I is divisible by $g(x)$. Use this to show that $I = \langle g(x) \rangle$.
Let $\epsilon > 0$ be given.
I is a subspace of X. Let $g(x) \in I$ have degree d, and note that this is the only polynomial in I with degree d.
Why? Well suppose $g(x) = ax + b$ and $g_2(x) = x$. Well then that would mean that $g(x) - ag_2(x) = b \in I$ which contradicts our statement that no nonzero polynomial of lower degree is in I. Thus the polynomial $g(x)$ is unique in the sense it is the only polynomial of degree d in I.
Now note that every polynomial in I is of degree *greater than* d.
Thus by the division algorithm, every $p(x) \in I$ with degree $m > d$ can be divided by $g(x)$ to produce a polynomial.

6. If $f(x)$ is a polynomial, and A is an $n \times n$ matrix, we can form the $n \times n$ matrix $f(A)$. For example, if $f(x) = x^2 + 2x + 3$, then $f(A) = A^2 + 2A + 3I_n$.

   (a) Show that if A is an $n \times n$ matrix, then
   $$\mathfrak{ann}(A) := \{f(x) \in \mathbb{F}[x] \mid f(A) = \mathbb{0}_{n\times n}\}$$
   is an ideal in $\mathbb{F}[x]$.
   Need to show that every polynomial in $\mathfrak{ann}(A)$ multiplied with a polynomial in $\mathbb{F}[x]$ produces a polynomial in $\mathfrak{ann}(A)$.
   With means to intimidate, this is easy to see since every $f(x) \in \mathfrak{ann}(A)$ multiplied with any $h(x) \in \mathbb{F}[x]$ will have the product $f(x)h(x) = \mathbb{0}_{n\times n} \in \mathfrak{ann}(A)$.

   (b) Find $\mathfrak{ann}(A)$, if
   $$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$
   We note that $A^n = \mathbb{0}_{2\times 2}$ for all $n \geq 2$. Thus any combination of $A^2 + A^3 + ... + A^n = \mathbb{0}_{2\times 2}$ for some n. Thus
   $$\mathfrak{ann}(A) = \{x^2, x^3, ..., x^n\}.$$
   It should be noted this could be an infinite set.