MATH 725 TAKE-HOME FINAL

MARIO L. GUTIERREZ ABED
PROF. J. LOUSTAU

The following set of exercises will result in a non-standard proof of the Cayley-Hamilton theorem. It is also a case of the statement: Once you have the Jordan form, then you can prove all theorems that pertain to matrices and linear transformations.

Note: It is standard to write a block diagonal matrix using direct sum notation, $A = A_1 \oplus \cdots \oplus A_n$. We refer to the size of A_i by n_i.

Note: We suppose throughout that the eigenvalues of all matrices lie in the scalar field F. Hence, all matrices have Jordan form.

Ex # 1) Let A and B be block diagonal $n \times n$ matrices with corresponding blocks of the same size. That is, with the notation just introduced, for each i, A_i and B_i have equal size, n_i. Prove that $C = AB$ is also block diagonal with $C_i = A_i B_i$. (This will follow from the basic formula for the ij entry of a product of matrices –the ij entry of a product $[s_{i,j}][t_{i,j}]$ is given by $\sum_k s_{i,k}t_{k,j}$.)

Proof. We write A and B as block diagonal $n \times n$ matrices,

$$A = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & A_n \end{pmatrix}_{\text{block}} \qquad \text{and} \qquad B = \begin{pmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & B_n \end{pmatrix}_{\text{block}}$$

Then the product $C = AB$ is defined as,

$$C_{i,j} = \sum_{k=1}^{n} A_{i,k} B_{k,j} = \begin{cases} 0 & \text{if } i \neq j, \\ A_i B_i & \text{if } i = j. \end{cases}$$

This gives us

$$C = \begin{pmatrix} A_1B_1 & 0 & \dots & 0 \\ 0 & A_2B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & A_nB_n \end{pmatrix}_{\text{block}},$$

which is block diagonal, as desired. □

Ex # 2) Prove that the determinant of an upper (lower) triangular matrix is the product of the diagonal entries.

Proof. Let A be an upper triangular matrix,

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{n,n} \end{pmatrix}.$$

Then the determinant of A can be defined using cofactors along its columns (or rows). WLOG, let us use cofactors long column 1:

$$\det(A) = |A| = a_{1,1} \begin{pmatrix} a_{2,2} & a_{2,3} & \dots & a_{2,n} \\ 0 & a_{3,3} & \dots & a_{3,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{n,n} \end{pmatrix} - a_{2,1} \begin{pmatrix} a_{1,2} & a_{1,3} & \dots & a_{1,n} \\ 0 & a_{3,3} & \dots & a_{3,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{n,n} \end{pmatrix} + \dots$$

$$+ (-1)^{i+1} a_{i,1} \begin{pmatrix} a_{1,2} & a_{1,3} & \dots & \dots & \dots & a_{1,n} \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & a_{i-1,i-1} & \dots & \vdots & a_{i-1,n} \\ \vdots & \ddots & 0 & a_{i+1,i+1} & \dots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & 0 & a_{n,n} \end{pmatrix} + \dots$$

$$+ (-1)^{n+1} a_{n,1} \begin{pmatrix} a_{1,2} & a_{1,3} & \dots & \dots & a_{1,n} \\ a_{2,2} & a_{2,3} & \dots & \dots & a_{2,n} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{n-1,n-1} & a_{n-1,n} \end{pmatrix}.$$

But since A is upper diagonal, we have that $a_{i,1} = 0$ for all $i \geq 2$. Hence the determinant is reduced to

$$\det(A) = a_{1,1} \begin{pmatrix} a_{2,2} & a_{2,3} & \dots & a_{2,n} \\ 0 & a_{3,3} & \dots & a_{3,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{n,n} \end{pmatrix}.$$

Then by recursively repeating the above procedure we find that

$$\det(A) = \prod_{i=1}^{n} a_{i,i},$$

which is the desired result. A similar procedure shows that this result also holds in the case when A is lower triangular. We just need to take cofactors along row 1, and then we get that $a_{1,j} = 0$ for all $j \geq 2$, from which follows that the determinant of A is the product of the diagonal elements as well. $\square$

Note: Let J be the Jordan form of A. Also recall that similar matrices have the same characteristic polynomial.

Ex # 3) Suppose the size of the i^{th} block J_i of J is n_i. The eigenvalues of A are α_i, where α_i is the diagonal entry of the Jordan block matrix J_i. Prove that the characteristic polynomial of A equals $\prod_i (\lambda - \alpha_i)^{n_i}$. (The issue here is the exponent of the monomials.)

Proof. As noted above, A is similar to J, so they must have the same characteristic polynomial. Note that since J is upper triangular, by the result on *Exercise 2*, it must be the case that the characteristic polynomial of J, $\det(\lambda I_n - J)$, must be equal to the product of the diagonal elements

$$(\lambda - \alpha_1)^{n_1}(\lambda - \alpha_2)^{n_2} \cdots (\lambda - \alpha_\ell)^{n_\ell},$$

where ℓ is the number of blocks in J. Hence, $\operatorname{char}(A) = \operatorname{char}(J) = \prod_i (\lambda - \alpha_i)^{n_i}$, as desired. □

Ex # 4) Prove that $(J - \alpha_i I_n)^{n_i}$ is block diagonal with the i^{th} block equal to zero.

Proof. To simplify notation, let $D = J - \alpha_i I_n$, such that

$$D_{s,t} = \begin{cases} \alpha_j - \alpha_i & \text{if } \ s = t, \\ 1 & \text{if } \ t = s + 1, \\ 0 & \text{otherwise.} \end{cases}$$

where j is the j^{th} block of J and $1 \leq s \leq n$, $1 \leq t \leq n$. Now, since the product of block diagonal matrices must be block diagonal, we have that since D is block diagonal, it follows that D^{n_i} is block diagonal as well. Then let B equal the i^{th} block of D, so that

$$B = \begin{pmatrix} \alpha_i - \alpha_i & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \alpha_i - \alpha_i \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}.$$

Similarly, D^{n_i} has an i^{th} block equal to B^{n_i}, such that,

$$B^{n_i}_{s,t} = \begin{cases} 1 & \text{for } \ B^{n_i}_{s,s+n_i} \\ 0 & \text{otherwise.} \end{cases}$$

Note that the index $s + n_i$ is greater than the size of i^{th} block, for all s, where $1 \leq s \leq n_i$. Therefore the i^{th} block of D^{n_i} is equal to zero, as desired. □

Ex # 5) Let $p(\lambda)$ be the characteristic polynomial of A. Prove that $p(J) = 0$.

Proof. Assuming there are ℓ blocks, we have that the characteristic polynomial of J is given by

$$p(J) = \prod_{i=1}^{\ell} (J - \alpha_i I_n)^{n_i}$$

Now, as we proved in the previous exercise, each i^{th} block of $(J - \alpha_i In)^{n_i}$ is equal to zero. Therefore the product of all the i^{th} blocks each raised to their corresponding n_i power is equal to zero. Thus, $p(J) = 0$, as desired. □

Ex # 6) (Cayley-Hamilton) Prove that $p(A) = 0$.

Proof. Let us start by letting $C = A - \alpha_i I_n$, so that

$$C = \begin{pmatrix} \alpha_i - \alpha_i & 0 & \dots & & & & & 0 \\ 0 & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & & \alpha_i - \alpha_i & & & & & \\ & & & \ddots & & & & \\ & & & & 0 & & & \\ & & & & & \ddots & & \\ & & & & & & \alpha_\ell - \alpha_i & \\ & & & & & & \ddots & 0 \\ 0 & \dots & & & & & 0 & \alpha_\ell - \alpha_i \end{pmatrix}.$$

Then we have

$$C_{s,t} = \begin{cases} \alpha_j - \alpha_i & \text{if } s = t, \\ 0 & \text{otherwise,} \end{cases}$$

where j is the j^{th} block of C. Thus,

$$C_{s,t}^{n_i} = \begin{cases} (\alpha_j - \alpha_i)^{n_i} & \text{if } s = t, j \neq i, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, we have that the i^{th} block of C^{n_i} is equal to zero. Then, following the same reasoning as in the previous exercise, we have

$$p(A) = \prod_{i=1}^{\ell} (A - \alpha_i I_n)^{n_i} = \prod_{i=1}^{\ell} C^{n_i} = 0. \qquad \square$$

Note: Let $p(\lambda) = \sum_{k=0}^{n} \beta_k \lambda^k$, the characteristic polynomial of A.

Ex # 7) Prove that $\beta_0 = (-1)^n \det(A)$ and it is the product of the eigenvalues (including multiplicity).

Proof. As before, assume that A has ℓ diagonal blocks. Then we have

$$p(\lambda) = \prod_{i=1}^{\ell} (\lambda - \alpha_i)^{n_i} = \prod_{i=1}^{\ell} (-1)^{n_i} (\alpha_i - \lambda)^{n_i} = \prod_{i=1}^{\ell} \sum_{k=0}^{n_i} (-1)^{n_i + k} \binom{n_i}{k} \lambda^k \alpha_i^{n_i - k}.$$

We can rewrite this in terms of summations as

$$(\clubsuit) \qquad p(\lambda) = \sum_{k_1=0}^{n_1} \cdots \sum_{k_\ell=0}^{n_\ell} (-1)^{n_1 + k_1} \binom{n_1}{k_1} \lambda^{k_1} \alpha_1^{n_1 - k_1} \cdots (-1)^{n_\ell + k_\ell} \binom{n_\ell}{k_\ell} \lambda^{k_\ell} \alpha_\ell^{n_\ell - k_\ell}.$$

In this expression the last term of the sum occurs when $k_1, \dots, k_\ell = n_1, \dots, n_\ell$. The first term appears when all $k_1, \dots, k_\ell = 0$. Also note that we have $\sum_{i=1}^{\ell} n_i = n$. This gives us

$$\begin{aligned} p(\lambda) &= \lambda^{n_1}\lambda^{n_2}\cdots\lambda^{n_\ell} + (-1)^{n_1+\cdots+n_\ell}\binom{n_1}{n_1}\cdots\binom{n_\ell}{n_\ell}(\lambda^{n_1-n_1}\alpha_1^{n_1}\cdots\lambda^{n_\ell-n_\ell}\alpha_\ell^{n_\ell}) \\ &= \lambda^n + (-1)^n(1\cdots 1)\prod_{i=1}^{\ell} 1\cdot\alpha_i^{n_i} \\ &= \lambda^n + (-1)^n\prod_{i=1}^{\ell}\alpha_i^{n_i}. \end{aligned}$$

Thus we have

$$\beta_0 = (-1)^n\prod_{i=1}^{\ell}\alpha_i^{n_i} = (-1)^n\det(A).$$

□

Ex # 8) Let $\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{i,i}$. (This is the sum of the diagonal entries of A, which is called the ***trace*** of A.)

- **a)** Prove that $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. (Same hint as for Exercise 1.)
- **b)** Prove that $\mathrm{Tr}(A) = \mathrm{Tr}(J)$.
- **c)** Prove that $\beta_{n-1} = \mathrm{Tr}(A)$, and it is the sum of the eigenvalues of A (including multiplicity).

Proof of ***a)***. Let $C = AB$ and $D = BA$, so that

$$C_{i,j} = \sum_{k=1}^{n} A_{i,k}B_{k,j} \qquad \text{and} \qquad D_{i,j} = \sum_{k=1}^{n} B_{i,k}A_{k,j}.$$

Then,

$$\begin{aligned} \mathrm{Tr}(C) = \sum_{i=1}^{n} C_{i,i} &= \sum_{i=1}^{n}\sum_{k=1}^{n} A_{i,k}B_{k,i} \\ &= \sum_{k=1}^{n}\sum_{i=1}^{n} B_{i,k}A_{k,i} \\ &= \sum_{i=1}^{n}\sum_{k=1}^{n} B_{i,k}A_{k,i} \\ &= \sum_{i=1}^{n} D_{i,i} = \mathrm{Tr}(D). \end{aligned}$$

□

Proof of ***b)***. Note that the diagonal elements of A are the α_i, for each of its ℓ blocks. Thus, we have

$$\mathrm{Tr}(A) = \sum_{i=1}^{\ell} n_i\alpha_i.$$

Now, for the i^{th} block of J, which we denote by J_i, we have the α_i as its diagonals. Thus,

$$\mathrm{Tr}(J_i) = \sum_{k=1}^{n_i} \alpha_i.$$

Hence, summing the traces of all the J_i, we get the desired result:

$$\mathrm{Tr}(J) = \sum_{i=1}^{\ell} \mathrm{Tr}(J_i) = \sum_{i=1}^{\ell} \sum_{k=1}^{n_i} \alpha_i = \sum_{i=1}^{\ell} n_i \alpha_i = \mathrm{Tr}(A). \qquad \square$$

*Proof of **c)**.* We showed in *Exercise 7* that $\beta_0 = (-1)^n \det(A)$. It follows that β_{n-1} corresponds to the λ with the $(n-1)^{st}$ exponent. Now we rearrange and reindex the terms of (♣) (see *Exercise 7*), so that λ^n appears when all $k_1, \ldots, k_\ell = 0$:

$$(♣♣) \qquad p(\lambda) = \sum_{k_1=0}^{n_1} \cdots \sum_{k_\ell=0}^{n_\ell} (-1)^{k_1} \binom{n_1}{k_1} \lambda^{n_1-k_1} \alpha_1^{k_1} \cdots (-1)^{k_\ell} \binom{n_\ell}{k_\ell} \lambda^{n_\ell-k_\ell} \alpha_\ell^{n_\ell}.$$

Now in order to get β_{n-1}, let $\sum_{i=1}^{\ell} k_i = 1$. That is, we take all possible permutations of values of $k_1, \ldots, k_\ell$, such that their sum is always 1. Then (♣♣) yields

$$p(\lambda) = \sum_{k_1=0}^{n_1} \cdots \sum_{k_\ell=0}^{n_\ell} (-1)^{k_1} \binom{n_1}{k_1} \alpha_1^{k_1} \cdots (-1)^{k_\ell} \binom{n_\ell}{k_\ell} \alpha_\ell^{n_\ell} \lambda^{n-(k_1+\cdots+k_\ell)}.$$

Thus,

$$\begin{aligned}
\beta_{n-1} &= \sum_{k_1+\cdots+k_\ell=1} \binom{n_1}{k_1} \alpha_1^{k_1} \cdots \binom{n_\ell}{k_\ell} \alpha_\ell^{k_\ell} \\
&= \binom{n_1}{1} \alpha_1 + \binom{n_2}{1} \alpha_2 + \cdots + \binom{n_\ell}{1} \alpha_\ell \\
&= \sum_{i=1}^{\ell} n_i \alpha_i = \mathrm{Tr}(A).
\end{aligned} \qquad \square$$

Ex # 9) Suppose that you have a Krylov process $\mathcal{K}_d(A, v)$, with $d = n$. What can you conclude about the Krylov process at v, at another vector w?

Solution. I'm not sure that I understood the question correctly but let's give it a try. Note that $\mathcal{K}_n(A, v)$ is the set of all linear combinations

$$(\dagger) \qquad a_0 v + a_1 A v + a_2 A^2 v + \cdots + a_{n-1} A^{n-1} v.$$

Given any coefficients $a_0, \ldots, a_{n-1}$, we can build a polynomial $q(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_{n-1} z^{n-1}$ of degree $n-1$ or less. Then the linear combination in (†) can be written more compactly as $q(A)v$. Thus, we have the following simple characterization of the Krylov subspace:

Let P_{n-1} denote the set of all polynomials of degree less than n. Then,

$$K_n(A, v) = \{q(A)v \mid q \in P_{n-1}\}.$$

Then the question of whether $K_n(A, v)$ contains good approximations to a given eigenvector w is therefore that of whether there are polynomials $q \in P_{n-1}$ such the $q(A)v \approx w$. $\square$