

University of Colorado Denver  
Department of Mathematical and Statistical Sciences  
Applied Linear Algebra Ph.D. Preliminary Exam  
June 10, 2011

Name: \_\_\_\_\_

**Exam Rules:**

- This is a closed book exam. Once the exam begins, you have 4 hours to do your best. Submit as many solutions as you can. All solutions will be graded and your final grade will be based on your six best solutions.
- Each problem is worth 20 points.
- Justify your solutions: **cite theorems that you use**, provide counter-examples for disproof, give explanations, and show calculations for numerical problems.
- If you are asked to prove a theorem, do not merely quote that theorem as your proof; instead, produce an independent proof.
- Begin each solution on a new page and use additional paper, if necessary.
- Write only on one side of paper.
- Write legibly using a dark pencil or pen.
- Ask the proctor if you have any questions.

Good luck!

1. _____	5. _____
2. _____	6. _____
3. _____	7. _____
4. _____	8. _____

Total \_\_\_\_\_

DO NOT TURN THE PAGE UNTIL TOLD TO DO SO.

**Applied Linear Algebra Preliminary Exam Committee:**  
Alexander Engau, Julien Langou (Chair), Richard Lundgren

1. (a) (8 pts) Let  $A$  be the  $n$ -by- $n$  matrix with all entries equal to one. Find the eigenvalues and corresponding eigenvectors of  $A$ .
- (b) (12 pts) Let  $B$  be the  $2n$ -by- $2n$  matrix

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

Find the eigenvalues and corresponding eigenvectors of  $B$ .

**Solution**

- (a)  $A$  has two distinct eigenvalues:  $n$  and  $0$ .  $A$  is actually diagonalizable. The geometric multiplicity of the eigenvalue  $n$  is 1, the geometric multiplicity of the eigenvalue  $0$  is  $n - 1$ . A basis of eigenvectors is for example:

$$\left( \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix} \right).$$

The first eigenvector given above is associated with the eigenvalue  $n$ . The following  $n - 1$  eigenvectors are associated with the eigenvalue  $0$ .

- (b)  $B$  has three distinct eigenvalues:  $n$ ,  $-n$ , and  $0$ .  $B$  is actually diagonalizable. The geometric multiplicity of the eigenvalue  $n$  is 1, the geometric multiplicity of the eigenvalue  $-n$  is 1, the geometric multiplicity of the eigenvalue  $0$  is  $2n - 2$ . A basis of eigenvectors is for example:

$$\left( \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ -1 \\ -1 \\ \vdots \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ -1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ -1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ -1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ -1 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ -1 \end{pmatrix} \right).$$

The first eigenvector given above is associated with the eigenvalue  $n$ . The second eigenvector given above is associated with the eigenvalue  $-n$ . The following  $2n - 2$  eigenvectors are associated with the eigenvalue 0.

2. Let  $\mathcal{P}_n$  be the vector space of all polynomials of degree at most  $n$  over  $\mathbb{R}$ . Define

$$\begin{aligned} T : \mathcal{P}_n &\longrightarrow \mathcal{P}_n \\ p(x) &\longmapsto p(x) = xp'(x) - p(x) \end{aligned}$$

- (a) (6 pts) Show that  $T$  is a linear transformation on  $\mathcal{P}_n$ .  
(b) (14 pts) Find  $\text{Null}(T)$  and  $\text{Range}(T)$ .

**Solution**

- (a) Let  $p$  and  $q$  be two polynomials of degree at most  $n$ , and let  $\lambda$  and  $\mu$  be two real numbers. Let  $x$  be a real number.  $(T(\lambda p + \mu q))(x) = x(\lambda p + \mu q)'(x) - (\lambda p + \mu q)(x) = \lambda(xp'(x) - p(x)) + \mu(xq'(x) - q(x)) = (\lambda T(p) + \mu T(q))(x)$ . So  $T(\lambda p + \mu q) = \lambda T(p) + \mu T(q)$ . So  $T$  is linear.
- (b) We consider the polynomial  $p(x)$  such that  $p(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0$ . A direct calculation gives

$$(T(p))(x) = a_n(n-1)x^n + a_{n-1}(n-2)x^{n-1} + \dots + a_2 x^2 + a_0.$$

We have

$$\text{Range}(T) = \text{Span}(1, x^2, x^3, \dots, x^n),$$

and

$$\text{Null}(T) = \text{Span}(x),$$

3. Let  $A$  be an  $m$ -by- $n$  real matrix and  $b \in \mathbb{R}^m$ . Show that exactly one of the following systems has a solution:

- (a)  $Ax = b$ ,
- (b)  $A^T y = 0$  and  $y^T b \neq 0$ .

### Solution

The name of the theorem is called the *Fredholm Alternative*. Two cases. First case: we assume that there is a solution for statement (a). (In this case, there exists  $x \in \mathbb{R}^n$  such that  $b = Ax$ .) We now prove that statement (b) has no solution. Let  $y \in \mathbb{R}^m$  such that  $A^T y = 0$ , multiplying on the left by  $x^T$ , we get  $x^T A^T y = (Ax)^T y = b^T y = 0$ . So that  $y^T b = 0$ . Statement (b) has no solution. Second case: we assume that there is no solution for statement (a). We need to prove that statement (b) has a solution. Let  $P$  be the orthogonal projection onto  $\text{Range}(A)$ . We claim that  $y = b - Pb = (I - P)b$  is a solution of (b). ( $y$  as defined is actually the orthogonal projection of  $b$  onto the orthogonal complement of  $\text{Range}(A)$ .) First of  $A^T y = 0$ . (Since  $y$  belongs to the orthogonal complement of  $A$ .) Now we want to prove that  $y^T b \neq 0$ . Let us assume that  $y^T b = 0$ . In this case,  $\|b\|^2 = b^T b = b^T (Pb + b - Pb) = b^T (Pb + y) = b^T Pb + b^T y = b^T Pb \leq \|b\| \cdot \|Pb\|$  (, the last inequality is Cauchy-Schwartz).  $b \neq 0$  because the case  $b = 0$  falls into Statement (a) has a solution, so, since  $b \neq 0$ , we find that,  $\|b\| \leq \|Pb\|$ , now  $P$  is an orthogonal projection so  $\|Pb\| \leq \|b\|$ , so that  $\|Pb\| = \|b\|$ , but since, by Pythagorean theorem,  $\|b\|^2 = \|Pb\|^2 + \|y\|^2$ , we get  $\|y\| = 0$ , this implies  $b = Pb$ , that is  $b \in \text{Range}(A)$ . This contradicts our assumption (, there is no solution for statement (a)). Therefore,  $y^T b \neq 0$ .

### Geometric Argument using the Separation Theorem for Convex Sets

As indicated in the above proof, this result has an intuitive geometric interpretation. Note that the range of  $A$  is a convex set, and that statement (a) is equivalent to vector  $b$  being in the set. By the separation theorem for convex sets, an element is either contained in the set (if statement (a) is true), or there exists a hyperplane  $y^T e = f$  for some nonzero normal vector  $y$  and some scalar  $f$  that separates set and element, such that  $y^T (Ax) \geq f$  for all  $x$  and  $y^T b < f$  (the inequalities could be reversed by alternating the sign of  $y$  and  $f$ ). Now, because  $y^T (Ax) \geq f$  for all  $x$  independent of the sign of  $x$ ,  $f$  must be zero and the inequality is actually an equality,  $y^T (Ax) = 0$  for all  $x$  and thus  $y^T A = 0$  with  $y^T b \neq 0$  by separation.

### Argument using the Duality Theorem in Linear Programming

In optimization theory, the above result is also known as a variant of Farkas Lemma and either used to prove, or derived from LP duality (if proven differently). From LP duality, a primal problem  $\min\{c^T x : Ax = b\}$  has an optimal solution if and only if its dual problem  $\max\{b^T y : A^T y = c\}$  has an optimal solution, in which case the

optimal values coincide, or the primal problem is infeasible if the dual problem is unbounded (and vice versa; note that this does not preclude that both primal and dual problem can also be infeasible, in general). Now, for  $c = 0$ , the dual problem is always feasible for  $y = 0$  with objective value 0. If 0 is the optimal value, then by duality there must be a feasible  $x$  for the primal, so  $Ax = 0$ , and there cannot be a feasible  $y$  for the dual (so  $A^T y = 0$ ) with objective  $y^T b \neq 0$  (because then there is  $\alpha$  greater or less than zero such that  $A^T(\alpha y) = 0$  but  $\alpha y^T b > 0$ , and 0 would not be optimal). Hence, in this case statement (a) is true and statement (b) is false. Following the same argument, if 0 is not the optimal value, then there is  $\alpha$  greater or less than zero such that  $A^T(\alpha y) = 0$  but  $\alpha y^T b > 0$  (so statement (b) is true), but then the dual problem is unbounded and the primal problem must be infeasible (so statement (a) is false).

4. Prove or disprove:

- (a) (7 pts) Let  $A$  be an  $n$ -by- $n$  matrix. If  $A^2 = 0$ , then the rank of  $A$  is at most 2.
- (b) (7 pts) Let  $T$  be a linear operator on a finite dimensional vector space  $V$  over  $\mathbb{R}$ . If  $T$  has no eigenvalues, then  $T$  is invertible.
- (c) (6 pts) Let  $A$  be Hermitian (self-adjoint). If  $A^2 = I$ , then  $A = I$  or  $-I$ .

**Solution**

- (a) **False.** Consider

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

(The matrix  $A$  has three Jordan block of size two associated with the eigenvalue 0.) Then  $A^2 = 0$  but  $\text{rank}(A) = 3$ .

- (b) **True.** We know that  $T$  is noninvertible implies that  $T$  has 0 as an eigenvalue. Therefore, if  $T$  has no eigenvalues,  $T$  does not have 0 as an eigenvalue (obviously), and so  $T$  is invertible.

Note: All said above is true in complex or real vector space, however in a complex vector space the statement  $T$  has no eigenvalues is false, therefore it is better to ask the problem in a real vector space.

- (c) **False.** Consider

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

5. If  $A$  is a nilpotent matrix, show that

- (a) (10 pts)  $I - A$  is invertible. Find  $(I - A)^{-1}$ .
- (b) (5 pts)  $I + A$  is invertible.
- (c) (5 pts)  $A$  is not diagonalizable when  $A$  is not the 0-matrix.

**Solution**

- (a) A matrix  $A$  is nilpotent if there exists an integer  $k$  such that  $A^k = 0$ . In this case, we note that

$$(I - A)(I + A + A^2 + \dots + A^{k-1}) = (I + A + A^2 + \dots + A^{k-1})(I - A) = I - A^k = I.$$

Therefore  $(I - A)$  is invertible and

$$(I - A)^{-1} = I + A + A^2 + \dots + A^{k-1}.$$

Note: this is related to the Taylor expansion of  $(1 + x)^{-1}$  for  $|x| < 1$ .

- (b) To find the inverse of  $I + A$ , we apply the previous formula to  $-A$  (, since  $A$  is nilpotent, so is  $-A$ ,) and obtain

$$(I + A)^{-1} = I - A + A^2 + \dots + (-1)^{k-1} A^{k-1}.$$

- (c) We will prove the contrapositive: If  $A$  is diagonalizable then  $A$  is the 0-matrix. Let us assume that  $A$  is diagonalizable. Then there exists a basis  $V$  such that  $A = VDV^{-1}$  where  $D$  contains the eigenvalue of  $A$  on the diagonal. Since  $A$  is nilpotent, its only eigenvalue is 0, so  $D = 0$ . We conclude that  $A = 0$ .



6. Let  $V$  be a complex vector space of finite dimension with  $\dim(V) = n < \infty$ . Let  $f \in \mathcal{L}(V)$  and  $g \in \mathcal{L}(V)$  such that  $fg = gf$ . Show that

- (a) (5 pts) if  $\lambda$  is an eigenvalue of  $f$ , then the eigenspace  $E_\lambda$  is invariant under  $g$ .
- (b) (5 pts)  $\text{Range}(f)$  and  $\text{Null}(f)$  are invariant under  $g$ .
- (c) (5 pts)  $f$  and  $g$  have at least one common eigenvector.
- (d) (5 pts) The matrix representations of  $f$  and  $g$  are both upper-triangular with respect to some basis.

**Solution**

(a) We recall that  $E_\lambda$  is defined as the set  $\{x \in V : f(x) = \lambda x\}$ . In other words,  $E_\lambda$  is the subspace made of 0 and all eigenvectors associated with the eigenvalue  $\lambda$ .

Let  $x \in E_\lambda$ , then  $f(x) = \lambda x$ , applying  $g$  on both sides gives  $g(f(x)) = g(\lambda x) = \lambda g(x)$ . Since  $f$  and  $g$  commute, we obtain that  $f(g(x)) = \lambda g(x)$ . So  $g(x) \in E_\lambda$ . So  $E_\lambda$  is invariant under  $g$ .

(b) Since  $\text{Null}(f) = E_0$ , (a) answers the question for the  $\text{Null}(f)$  case:  $\text{Null}(f)$  is invariant under  $g$ . Let  $x \in \text{Range}(f)$ , then there exists  $y$  in  $V$  such that  $x = f(y)$ . Applying  $g$  on both sides gives  $g(x) = g(f(y))$ . Commuting  $f$  and  $g$  gives  $g(x) = f(g(y))$ . So  $g(x) \in \text{Range}(f)$ . So  $\text{Range}(f)$  is invariant under  $g$ .

(c) Let  $\lambda$  be an eigenvalue of  $f$ . By (a), we know that  $E_\lambda$  is invariant under  $g$ . We consider now the linear mapping  $g|_{E_\lambda} : E_\lambda \rightarrow E_\lambda$ . Since,  $E_\lambda$  is a *complex* vector space, any  $f \in \mathcal{L}(E_\lambda)$  has at least one eigenvalue. So  $g|_{E_\lambda}$  has an eigenvalue with an associated eigenvector, say  $y$ . This vector  $y$  is eigenvector of  $g$  (since it is an eigenvector of its restriction  $g|_{E_\lambda}$ ) but is also an eigenvector for  $f$  (since it belongs to  $E_\lambda$ ).  $f$  and  $g$  have a common eigenvector.

(d) This is an inductive process. From (c) and the commutativity of  $f$  and  $g$ , we obtain there exists a common eigenvector  $y_1$  for  $f$  and  $g$ . Hence, there are corresponding eigenvalues  $\lambda_1$  and  $\mu_1$  such that

$$f(y_1) = \lambda_1 y_1 \text{ and } g(y_1) = \mu_1 y_1.$$

We complete  $\text{span}(y_1)$  with a subspace  $\mathcal{U}_2$  of dimension  $n - 1$  such that  $\text{Span}(y_1) \oplus \mathcal{U}_2 = V$ . The matrix of  $f$  and  $g$  in a basis made of  $y_1$  and basis of  $\mathcal{U}_2$  looks like:

$$\begin{pmatrix} \star & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \\ 0 & \star & \star & \star \end{pmatrix}$$

We will use  $P_2$  the oblique projection onto  $\mathcal{U}_2$  parallel to  $y_1$ .  $P_2$  is defined as followed. Let  $x \in V$ , since  $\text{Span}(y_1) \oplus \mathcal{U}_2 = V$ , there exist unique  $x_1 \in$

Span( $y_1$ ) and unique  $x_2 \in \mathcal{U}_2$ , such that  $x = x_1 + x_2$ .  $P_2(x)$  is defined as  $P_2(x) = x_2$ . We now define  $f_2 : \mathcal{U}_2 \rightarrow \mathcal{U}_2$  as the composition of  $f$  followed by the projection onto  $\mathcal{U}_2$  parallel to  $y_1$  and  $g_2 : \mathcal{U}_2 \rightarrow \mathcal{U}_2$  as the composition of  $g$  followed by the projection onto  $\mathcal{U}_2$  parallel to  $y_1$ . In other words:

$$f_2 = P_2 f \text{ and } g_2 = P_2 g.$$

We have that  $f_2 \in \mathcal{L}(\mathcal{U}_2)$ ,  $g_2 \in \mathcal{L}(\mathcal{U}_2)$ , and  $f_2$  and  $g_2$  commute. So, (c) gives us  $y_2$  a common eigenvector for  $f_2$  and  $g_2$ . There exists  $\lambda_2$  and  $\mu_2$  such that

$$f_2(y_2) = \lambda_2 y_2 \text{ and } g_2(y_2) = \mu_2 y_2. \quad (1)$$

So that

$$P_2(f(y_2)) = \lambda_2 y_2 \text{ and } P_2(g(y_2)) = \mu_2 y_2.$$

Now, since  $P_2$  is the oblique projection onto  $\mathcal{U}_2$  parallel to  $y_1$ , for all  $x$ , there exists  $\alpha$  such that  $x = \alpha y_1 + P_2 x$ . Applying this to  $f(y_2)$  and  $g(y_2)$ , we see that there exist  $s_{12}$  and  $t_{12}$  such that

$$f(y_2) = s_{12} y_1 + P_2(f(y_2)) \text{ and } g(y_2) = t_{12} y_1 + P_2(g(y_2)).$$

Combining Eq (??) with the previous relation gives

$$f(y_2) = s_{12} y_1 + \lambda_2 y_2 \text{ and } g(y_2) = t_{12} y_1 + \mu_2 y_2.$$

We complete span( $y_1, y_2$ ) with a subspace  $\mathcal{U}_3$  of dimension  $n - 2$  such that  $\text{Span}(y_1, y_2) \oplus \mathcal{U}_3 = V$ . The matrix of  $f$  and  $g$  in a basis made of  $(y_1, y_2)$  and basis of  $\mathcal{U}_3$  looks like:

$$\begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & * & * \end{pmatrix}$$

We continue the process.

Note: if we would have an inner product, there would be no difficulty in finding an *orthonormal* basis with respect to which the matrix representations of  $f$  and  $g$  are both upper-triangular. The complement subspaces  $\mathcal{U}_1, \mathcal{U}_2, \dots$  need to be taken as  $\mathcal{U}_2 = y_1^\perp, \mathcal{U}_3 = y_2^\perp, \dots$

7. (a) (5 pts) Let  $A$  be a 5-by-5 matrix. Suppose that you know that  $\text{rank}(A^2) < 5$ . What can you say about  $\text{rank}(A)$ ?

(b) (10 pts) Write down

$$\det \begin{pmatrix} x & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{pmatrix}$$

as a polynomial in  $x$ , either factored or expanded.

(c) (5 pts) What are the singular values of a  $1 \times n$  matrix? What is the pseudo-inverse of a  $1 \times n$  matrix?

### Solution

(a)  $\text{rank}(A)$  is 0, 1, 2, 3, or 4. Since  $A$  is a 5-by-5 matrix, then  $\text{rank}(A) \leq 5$ . If  $\text{rank}(A) = 5$ ,  $A$  is full rank, and so  $\text{rank}(A^2) = 5$ . Not possible. Now all values 0, 1, 2, 3, or 4 are possible. For 3, we can for example consider the matrix  $A$  with all zeros except 3 ones on the diagonal. Replace 3 by 0, 1, 2, or 4 in the previous sentence to get all the cases.

(b) **Answer 1:** We define  $A$  as

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

The polynomial in  $x$  we are looking for is nothing else as the characteristic polynomial of  $A$  up to the sign of  $x$ .  $A$  has two distinct eigenvalues: 3 and  $-1$ .  $A$  is actually diagonalizable. The geometric multiplicity of the eigenvalue 3 is 1, the geometric multiplicity of the eigenvalue  $-1$  is 3. We conclude that the characteristic polynomial of  $A$  is

$$\chi_A(\lambda) = (\lambda + 1)^3(\lambda - 3),$$

(there is a  $+$  sign on  $\chi_A$  since 4 is even,) and so replacing  $\lambda$  by  $-x$  gives us

$$\chi_A(\lambda) = (x - 1)^3(x + 3).$$

A basis of eigenvectors is for example:

$$\left( \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ -1 \\ -1 \end{pmatrix}, \dots, \begin{pmatrix} -1 \\ -1 \\ 3 \\ -1 \end{pmatrix} \right).$$

The first eigenvector given above is associated with the eigenvalue 3. The following 3 eigenvectors are associated with the eigenvalue  $-1$ .

**Answer 2:** A direct computation is as easy:

$$\begin{aligned}
 \begin{vmatrix} x & 1 & 1 & 1 \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ 1 & 1 & 1 & x \end{vmatrix} &= - \begin{vmatrix} 1 & 1 & 1 & x \\ 1 & x & 1 & 1 \\ 1 & 1 & x & 1 \\ x & 1 & 1 & 1 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 1 & x \\ 0 & x-1 & 0 & 1-x \\ 0 & 0 & x-1 & 1-x \\ 0 & 1-x & 1-x & 1-x^2 \end{vmatrix} \\
 &= - \begin{vmatrix} 1 & 1 & 1 & x \\ 0 & x-1 & 0 & 1-x \\ 0 & 0 & x-1 & 1-x \\ 0 & 0 & 1-x & 2-x-x^2 \end{vmatrix} = - \begin{vmatrix} 1 & 1 & 1 & x \\ 0 & x-1 & 0 & 1-x \\ 0 & 0 & x-1 & 1-x \\ 0 & 0 & 0 & 3-2x-x^2 \end{vmatrix} \\
 &= (x-1)^3(3+x)
 \end{aligned}$$

- (c) Let  $x$  be 1-by- $n$  matrix. Then there is one singular value and its value is  $\sigma = \|x\|_2 = \sqrt{\sum_i x_i^2}$ . The singular value decomposition of the  $x$  is  $x = u\sigma v^T$ , where  $u$  is simply the scalar 1,  $\sigma$  is as previously given ( $\sigma = \|x\|_2$ ), and  $v = \frac{1}{\|x\|_2}x^T$ . The pseudoinverse is given by  $x^\dagger = v(\sigma)^{-1}u^T$  for  $\sigma \neq 0$ , and  $x^\dagger = 0$  for  $\sigma = 0$ . Answer: if  $x = 0$ ,  $x^\dagger = 0$ , if  $x \neq 0$ ,  $x^\dagger = \frac{1}{\|x\|_2^2}x^T$

Note: We can check that  $xx^\dagger = \frac{1}{\|x\|_2^2}xx^T$  is the orthogonal projection on  $x$ .

8. The Cayley-Hamilton theorem states that any square matrix satisfies its own characteristic equation. Prove it, following the steps below. Do not deal with the real arithmetic case, work only in complex arithmetic.

- (a) (7 pts) Prove that the theorem holds for square matrices that may be diagonalized.
- (b) (7 pts) Prove that the theorem holds for Jordan blocks, i.e., matrices of the form  $\lambda I + J$ , where  $\lambda \in \mathbb{C}$ ,  $I$  is the identity matrix and  $J$  is the matrix with zeros everywhere, except immediately above the diagonal, where it has 1's.
- (c) (6 pts) Prove the theorem for all square matrices.

**Solution**

Let  $\chi$  be the characteristic polynomial of  $A$ , we want to prove that

$$\chi(A) = 0.$$

Let  $n$  be the order of  $A$ . Let us assume that  $A$  has  $k$  distinct eigenvalues in  $\mathbb{C}$ :  $\lambda_1, \dots, \lambda_k$ . Each eigenvalue of  $A$  is a root of the characteristic polynomial of  $A$  so that the characteristic polynomial of  $A$  writes  $\chi(x) = (\lambda_1 - x)^{m_1} \dots (\lambda_k - x)^{m_k}$ , where  $m_i$  is an integer representing the algebraic multiplicity of  $\lambda_i$ . We have that  $m_i \geq 1$  and  $\sum m_i = n$ .

- (a) Let assume that  $A$  is diagonalizable matrix. There exists an invertible matrix  $V$  (basis of eigenvectors) and a diagonal matrix such that  $A = VDV^{-1}$ . Now

$$\chi(A) = \chi(VDV^{-1}) = V\chi(D)V^{-1} = V \begin{pmatrix} \chi(\lambda_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \chi(\lambda_k) \end{pmatrix} V^{-1}.$$

Since, for all  $i$ ,  $\chi(\lambda_i) = 0$ , then

$$\chi(A) = 0.$$

- (b) Let  $A$  be a Jordan block. So there exists  $\lambda \in \mathbb{C}$ , so that

$$A = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

The characteristic polynomial of  $A$  is

$$\chi(x) = (\lambda - x)^n.$$

Now

$$\chi(A) = (\lambda I - A)^n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}^n = 0.$$

- (c) We first reduce  $A$  to Jordan form. (This is possible for any matrix.) We know that there exists  $V$  an invertible matrix and  $J$  a block diagonal matrix made of Jordan blocks on the diagonal such that  $A = VJV^{-1}$ . We assume that  $A$  has  $\ell$  Jordan blocks:  $J_1, \dots, J_\ell$  of dimension  $n_1, \dots, n_\ell$ . In this case the characteristic polynomial of  $A$  writes  $\chi(x) = (\lambda_1 - x)^{n_1} \dots (\lambda_\ell - x)^{n_\ell}$ . (Note: the indexing of the eigenvalue  $\lambda$  here is different than the one in part (a). Here we allow for repeat.  $\ell$  is the number of Jordan blocks.) In this case, we have

$$\chi(A) = \chi(VJV^{-1}) = V\chi(J)V^{-1} = V \begin{pmatrix} \chi(J_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \chi(J_k) \end{pmatrix} V^{-1}.$$

We know by part (b) that, for all  $i = 1, \dots, \ell$ ,  $((\lambda_i - x)^{n_i})(J_i) = 0$ , so a fortiori,  $\chi(J_i) = 0$ . Finally,

$$\chi(A) = 0.$$