University of Colorado Denver
Department of Mathematical and Statistical Sciences
Applied Linear Algebra Ph.D. Preliminary Exam Solutions
August 9, 2021

Name: ____

Exam Rules:

- This exam is being administered remotely using Zoom. During the exam, you must be logged in to the assigned Zoom meeting with your your camera on, and you must be visible in the camera. Your microphone can be muted, but please leave Zoom audio ON so that the proctor can speak to you if needed.
- If you have any questions during the exam, please contact the exam proctor using the zoom chat, or call the proctor at the number given below.

  Committee Member Contact Information:

Name	Phone	email	Proctoring times
Julien Langou		julien.langou@ucdenver.edu	
Yaning Liu		yaning.liu@ucdenver.edu	
Jan Mandel	xxx-xxx-xxxx	jan.mandel@ucdenver.edu	10:00am-2:00pm

- If you need to be out of view of the camera for any reason (bathroom breaks, etc.), please let the proctor know by posting a message on the zoom chat.
- If you would like to work with a paper copy of the exam, please print it as soon as you receive it and inform the proctor that you are doing so before the exam begins. If the printer is in another room, let the proctor know.
- You may read the exam as soon as you receive it, but you may not start writing (even your name) until authorized to start writing.
- This is a closed book exam. You may not use any external aids during the exam, such as:
  - communicating with anyone other than the exam proctor (through text messages or emails, for example);
  - consulting the internet, textbooks, notes, solutions of previous exams, etc;
  - using calculators or mathematical software.
- You may use a tablet PC (such as an iPad or Microsoft Surface) to write your solutions. Alternatively, you can write your solutions on paper.
- Please begin each problem on a new page, and write the problem number and page number at the top of each page. (For example, 6-1, 6-2, 6-3 for pages 1, 2 and 3 of problem 6). If you are writing on paper, write only on one side of the paper.
- The exam will end 4 hours after it begins. At the conclusion of the exam, please email a copy of your solutions to the exam proctor. Do not leave until the proctor acknowledges that your solutions have been successfully received.

  - Your solutions need to be in a single .pdf file with the pages in the correct order. The .pdf file needs to be of good enough quality for easy grading.
  - If you cannot create a good quality .pdf file quickly, you may instead submit an imperfect scan, or even pictures of your exam, and then take more time to prepare and submit a good quality .pdf version. We will grade the better version but use the first submission to check that nothing was added or changed between versions.
  - Do not submit your scratch work.

- There are 8 total problems. Do all 4 problems in the first part (problems 1 to 4), and pick two problems in the second part (problems 5 to 8). Do not submit more than two solved problems from the second part. If you do, only the first two attempted problems will be graded. Each problem is worth 20 points.
- Do not submit multiple alternative solutions to any problem; if you do, only the first solution will be graded.
- 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.
- <u>Notation</u>: Throughout the exam, $\mathbb{R}$ and $\mathbb{C}$ denote the sets of real and complex numbers, respectively. $\mathbb{F}$ denotes either $\mathbb{R}$ or $\mathbb{C}$. $\mathbb{F}^n$ and $\mathbb{F}^{n,n}$ are the vector spaces of n-tuples and $n \times n$ matrices, respectively, over the field $\mathbb{F}$. $\mathcal{L}(V)$ denotes the set of linear operators on the vector space V. T^* is the adjoint of the operator T and λ^* is the complex conjugate of the scalar λ. In an inner product space V, $U^\perp$ denotes the orthogonal complement of the subspace U.
- If you are confused or stuck on a problem, either ask a question or move on to another problem.

Problem	Points	Score		Problem	Points	Score
1.	20			5.	20	
2.	20			6.	20	
3.	20			7.	20	
4.	20			8.	20	
				Total	120	

Applied Linear Algebra Preliminary Exam Committee:
Julien Langou, Yaning Liu (Chair), Jan Mandel

Before starting the exam, please sign and date the following honor statement.

<u>Honor statement:</u> I attest that I will not cheat and will not attempt to cheat. I attest that I will not communicate with anyone while taking the exam, I will not look at notes, textbooks, previous solutions of exams, cheat sheets, etc. I attest that I will not go on the web to find solutions. If I in some way receive information during the exam that might help with the exam, I will let the proctor know.

Part I. Work **all** of problems 1 through 4.

Problem 1. Let A be an $m \times n$ matrix over a field $\mathbb{F}$. Prove the following statements:

(a) $\text{null}A = \{\mathbf{0}\}$ if and only if the columns of A are linearly independent.

(b) If the columns of A are linearly independent, are the rows of A necessarily linearly independent?

(c) If $m < n$, then $\text{null}A \neq \{\mathbf{0}\}$.

(d) If $A = BC$, where B is $m \times m$ and C is $m \times n$, and if B is nonsingular, then $\text{null}A = \text{null}C$.

(e) Let A^* be the conjugate transpose of A. Then $\text{null}(A^*A) = \text{null}A$.

Solution

(a) Let $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n]$, where $\boldsymbol{a}_i \in \mathbf{F}^m$ is the ith column of A. Then $\text{null}A = \{\mathbf{0}\}$ if and only if $A\boldsymbol{x} = \mathbf{0}$ has the unique solution $\boldsymbol{x} = \mathbf{0}$, i.e.,

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \cdots + c_x\boldsymbol{a}_n = \mathbf{0}$$

if and only if $c_1 = c_2 = \cdots = c_n = 0$

(b) If the columns of A are linearly independent, then $m \geq n$. If $m > n$, then the rows of A are linearly dependent. If $m = n$, then the rows of A are linearly independent.

If the columns of A are linearly independent, then the rows of A are not necessarily linearly independent. For example, the matrix

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

have linearly dependent rows and linearly independent columns.

(c) If $A\boldsymbol{x} = \mathbf{0}$, then $\text{rank}A \leq m < n$. So by the rank-nullity theorem, we have $\dim \text{null}A = n - \text{rank}A > 0$.

(d) If $A\boldsymbol{x} = \mathbf{0}$, then $BC\boldsymbol{x} = \mathbf{0}$. If B is invertible, then $B^{-1}BC\boldsymbol{x} = C\boldsymbol{x} = \mathbf{0}$. So $\text{null}A \subseteq \text{null}C$. On the other hand, if $C\boldsymbol{x} = \mathbf{0}$, $BC\boldsymbol{x} = A\boldsymbol{x} = \mathbf{0}$, so $\text{null}C \subseteq \text{null}A$. So $\text{null}A = \text{null}C$.

(e) If $A\boldsymbol{x} = \mathbf{0}$, then $A^*A\boldsymbol{x} = \mathbf{0}$. So $\text{null}A \subseteq \text{null}(A^*A)$. Since $\text{rank}(A^*A) = \text{rank}A$, we have $\dim \text{null}A = \dim \text{null}(A^*A)$, by the rank-nullity theorem. It follows that $\text{null}A = \text{null}(A^*A)$.

Another solution: Let $\boldsymbol{x} \in \text{null}A$. Then $A\boldsymbol{x} = \mathbf{0}$, it follows that $A^*A\boldsymbol{x} = \mathbf{0}$, so $\boldsymbol{x} \in \text{null}(A^*A)$. Now suppose that $\boldsymbol{x} \in \text{null}(A^*A)$, thus $A^*A\boldsymbol{x} = \mathbf{0}$. Then $0 = \boldsymbol{x}^*A^*A\boldsymbol{x} = (A\boldsymbol{x})^*(A\boldsymbol{x})$ thus $A\boldsymbol{x} = \mathbf{0}$. So $\boldsymbol{x} \in \text{null}A$.

Problem 2. Let V, W be finite dimensional real inner product spaces and let $T : V \to W$ be a linear map. Fix $\boldsymbol{w} \in W$. Show that the elements $\boldsymbol{v} \in V$ for which the Euclidean norm $||T\boldsymbol{v} - \boldsymbol{w}||$ is minimal are exactly the solutions to the equation $T^*T\boldsymbol{v} = T^*\boldsymbol{w}$.

Solution Note that $\text{range}\, T$ is a subspace of W. For any $\boldsymbol{w} \in W$, let $\boldsymbol{w} = \boldsymbol{w}_1 + \boldsymbol{w}_2$ be the orthogonal project of $\boldsymbol{w}$ onto $\text{range}\, T$, i.e., $\boldsymbol{w}_1 \in \text{range}\, T$ and $\boldsymbol{w}_2 \in (\text{range}\, T)^{\perp}$. So the minimizers of $||T\boldsymbol{v} - \boldsymbol{w}||$ are those $\boldsymbol{v} \in V$ such that $T\boldsymbol{v} = \boldsymbol{w}_1$. Now we show that these satisfy the normal equation.

Suppose $\boldsymbol{v} \in V$ is such that $T\boldsymbol{v} = \boldsymbol{w}_1$. Then $\boldsymbol{w}_2 = \boldsymbol{w} - \boldsymbol{w}_1 = \boldsymbol{w} - T\boldsymbol{v}$ is orthogonal to $\text{range}\, T$. So for all $\boldsymbol{x} \in V$, it follows that

$$\langle T\boldsymbol{x}, T\boldsymbol{v} - \boldsymbol{w}\rangle = 0 \Rightarrow \langle \boldsymbol{x}, T^*(T\boldsymbol{v} - \boldsymbol{w})\rangle = 0$$

So $T^*(T\boldsymbol{v} - \boldsymbol{w})$ is orthogonal to V and thus $T^*(T\boldsymbol{v} - \boldsymbol{w}) = \mathbf{0}$ and so $T^*T\boldsymbol{v} = T^*\boldsymbol{w}$.

Conversely, let $T^*T\boldsymbol{v} = T^*\boldsymbol{w}$. It suffices to show $T\boldsymbol{v} = \boldsymbol{w}_1$. Consider

$$\begin{aligned}
\langle T\boldsymbol{v} - \boldsymbol{w}_1, T\boldsymbol{v} - \boldsymbol{w}_1\rangle &= \langle T\boldsymbol{v} - \boldsymbol{w}_1, T\boldsymbol{v} - (\boldsymbol{w} - \boldsymbol{w}_2)\rangle \\
&= \langle T\boldsymbol{v} - \boldsymbol{w}_1, (T\boldsymbol{v} - \boldsymbol{w}) + \boldsymbol{w}_2\rangle \\
&= \langle T\boldsymbol{v}, T\boldsymbol{v} - \boldsymbol{w}\rangle + \langle T\boldsymbol{v}, \boldsymbol{w}_2\rangle - \langle \boldsymbol{w}_1, T\boldsymbol{v} - \boldsymbol{w}\rangle - \langle \boldsymbol{w}_1, \boldsymbol{w}_2\rangle \\
&= \langle \boldsymbol{v}, T^*(T\boldsymbol{v} - \boldsymbol{w})\rangle + \langle T\boldsymbol{v}, \boldsymbol{w}_2\rangle - \langle \boldsymbol{w}_1, T\boldsymbol{v} - \boldsymbol{w}\rangle - \langle \boldsymbol{w}_1, \boldsymbol{w}_2\rangle \\
&= -\langle \boldsymbol{w}_1, T\boldsymbol{v} - \boldsymbol{w}\rangle
\end{aligned}$$

Note that $\boldsymbol{w}_1 \in \text{range}\, T$, so $\boldsymbol{w}_1 = T\boldsymbol{x}$ for some $\boldsymbol{x} \in V$. So previous equation can be rewritten as

$$\langle T\boldsymbol{v} - \boldsymbol{w}_1, T\boldsymbol{v} - \boldsymbol{w}_1\rangle = -\langle T\boldsymbol{x}, T\boldsymbol{v} - \boldsymbol{w}\rangle = -\langle \boldsymbol{x}, T^*(\boldsymbol{v} - \boldsymbol{w})\rangle = 0$$

So we have $||T\boldsymbol{v} - \boldsymbol{w}_1|| = 0$ and thus $T\boldsymbol{v} = \boldsymbol{w}_1$. So $\boldsymbol{v}$ is a minimizer of $||T\boldsymbol{v} - \boldsymbol{w}||$.

Another solution For $\boldsymbol{h} \in V$, define

$$F(\boldsymbol{h}) = ||T(\boldsymbol{v} + \boldsymbol{h}) - \boldsymbol{w}||^2 = \langle T(\boldsymbol{v} + \boldsymbol{h}) - \boldsymbol{w}, T(\boldsymbol{v} + \boldsymbol{h}) - \boldsymbol{w}\rangle.$$

Then, $||T\boldsymbol{v} - \boldsymbol{w}||$ is minimal if and only if the minimum of $F(\boldsymbol{h})$ is attained at $\boldsymbol{h} = \boldsymbol{0}$. By collecting the terms with $\boldsymbol{h}$ and using the definition of T^*,

$$\begin{aligned} F(\boldsymbol{h}) =& \langle T\boldsymbol{h}, T\boldsymbol{h}\rangle + 2\langle T\boldsymbol{h}, T\boldsymbol{v} - \boldsymbol{w}\rangle + c(\boldsymbol{v}, \boldsymbol{w}) \\ =& \langle T\boldsymbol{h}, T\boldsymbol{h}\rangle + 2\langle \boldsymbol{h}, T^*T\boldsymbol{v} - T^*\boldsymbol{w}\rangle + c(\boldsymbol{v}, \boldsymbol{w}), \end{aligned}$$

where $c(\boldsymbol{v}, \boldsymbol{w})$ does not depend on $\boldsymbol{h}$.

If $T^*T\boldsymbol{v} - T^*\boldsymbol{w} = 0$, then $F(\boldsymbol{h}) = \langle T\boldsymbol{h}, T\boldsymbol{h}\rangle + c(\boldsymbol{v}, \boldsymbol{w}) \geq c(\boldsymbol{v}, \boldsymbol{w}) = F(\boldsymbol{0})$, for any $\boldsymbol{h}$, so the minimum of $F(\boldsymbol{h})$ is attained at $\boldsymbol{h} = 0$.

If $T^*T\boldsymbol{v} - T^*\boldsymbol{w} \neq 0$, define $\boldsymbol{h}_t = t\boldsymbol{h}_0$, $\boldsymbol{h}_0 = T^*T\boldsymbol{v} - T^*\boldsymbol{w}$, $t \in \mathbb{R}$, and $f(t) = F(t\boldsymbol{h}_0)$. Then, $f(t) = t^2\langle T\boldsymbol{h}_0, T\boldsymbol{h}_0\rangle + 2t\langle \boldsymbol{h}_0, \boldsymbol{h}_0\rangle + c(\boldsymbol{v}, \boldsymbol{w})$. Then, the derivative

$$f'(0) = 2\langle T^*T\boldsymbol{v} - T^*\boldsymbol{w}, T^*T\boldsymbol{v} - T^*\boldsymbol{w}\rangle\rangle 0$$

so for some $t\langle 0$, we have $f(t)\langle 0$ and thus $F(t\boldsymbol{h}_0) = f(t)\langle 0 = F(\boldsymbol{0})$, and $F(\boldsymbol{h})$ does not have minimum at $\boldsymbol{h} = 0$.

Problem 3.

For each of the following 4 statements, give either a counterexample or a reason why it is true.

1. For every real matrix A there is a real matrix B with $B^{-1}AB$ diagonal.
2. For every symmetric real matrix A there is a real matrix B with $B^{-1}AB$ diagonal.
3. For every complex matrix A there is a complex matrix B with $B^{-1}AB$ diagonal.
4. For every symmetric complex matrix A there is a complex matrix B with $B^{-1}AB$ diagonal.

Solution

1. False. Not all real matrices are diagonalizable. For example

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

   is not diagonalizable (in real arithmetic).

2. True. Per real spectral theorem, a symmetric real matrix is diagonalizable with real eigenvalues and real eigenvectors.

3. False. Not all complex matrices are diagonalizable. For example

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

is not diagonalizable (in complex arithmetic).

4. False. Not all symmetric complex matrices are diagonalizable. For example

$$A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$$

is symmetric and is not diagonalizable (in complex arithmetic). Why is A not diagonalizable? Its trace is zero, and its determinant is zero. So the characteristic polynomial is λ^2. So 0 is the only eigenvalue of A. The only 2x2 diagonalizable matrix that has zero as its only eigenvalue is the 2x2 zero matrix. Clearly A is not the zero matrix, so A is not diagonalizable.

Note: Complex selfadjoint matrices are diagonalizable. But for a complex matrix, being selfadjoint and symmetric are not the same.

Problem 4. Suppose $T \in \mathcal{L}(V)$ and $(T - 2I)(T - 3I)(T - 4I) = 0$. Suppose λ is an eigenvalue of T. Prove that $\lambda = 2$ or $\lambda = 3$ or $\lambda = 4$.

Solution Since $(T - 2I)(T - 3I)(T - 4I) = 0$, the operator $(T - 2I)(T - 3I)(T - 4I)$ is not injective. Since the composition of injective maps is injective, at least one of the operators $T - 2I$, $T - 3I$, $T - 4I$ is not injective, thus $\lambda = 2$ or $\lambda = 3$ or $\lambda = 4$ is eigenvalue of T.

Another solution. Denote $p(z) = (z - 2)(z - 3)(z - 4)$. Since $p(T) = 0$, the minimal polynomial q of T divides p. Since eigenvalues of T are zeros of the minimal polynomial, the monomial $z - \lambda$ divides the minimal polynomial q, which divides the polynomial p, thus $z - \lambda$ divides $p(z)$, so $z - \lambda$ is one of the factors $z - 2$, $z - 3$, or $z - 4$.

Quoting and applying the spectral mapping theorem is also acceptable.

Part II. Work **two** of problems 5 through 8.

Problem 5. Consider vectors in $\mathbb{C}^n$ as columns, i.e. matrices of size n by 1. Then for two vectors $\boldsymbol{u}, \boldsymbol{v}$, the product $\boldsymbol{u}\boldsymbol{v}^*$ is an n by n matrix. It is called a "rank-one matrix" when $\boldsymbol{u}$ and $\boldsymbol{v}$ are both nonzero.

(a) Suppose that $A = A^* \in \mathbb{C}^{n,n}$. Show that A is a linear combination of rank-one matrices $\boldsymbol{u}_k\boldsymbol{u}_k^*$, $k = 1, \dots, n$, where $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ is an orthonormal basis of $\mathbb{C}^n$.

(b) Suppose that $A \in \mathbb{C}^{n,n}$. Show that A is a linear combination of rank-one matrices of the form $\boldsymbol{f}_k\boldsymbol{e}_k^*$, $k = 1, \dots, n$, where $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ and $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$ are orthonormal bases.

Solution (a) Since $A = A^*$, $\mathbb{C}^n$ has a basis consisting of orthonormal eigenvectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ of A,

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i$$

Let $\boldsymbol{v} \in \mathbb{C}^n$. Then $\boldsymbol{v}$ has the expansion

$$\boldsymbol{v} = \sum_{i=1}^{n} \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle \boldsymbol{u}_i,$$

and

$$A\boldsymbol{v} = \sum_{i=1}^{n} \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle \lambda \boldsymbol{u}_i = \sum_{i=1}^{n} \lambda \boldsymbol{u}_i \langle \boldsymbol{v}, \boldsymbol{u}_i \rangle = \sum_{i=1}^{n} \lambda \boldsymbol{u}_i \boldsymbol{u}_i^* \boldsymbol{v} = \left(\sum_{i=1}^{n} \lambda \boldsymbol{u}_i \boldsymbol{u}_i^* \right) \boldsymbol{v}.$$

Since $\boldsymbol{v} \in \mathbb{C}^n$ was arbirtrary,

$$A = \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^*.$$

(b) The SVD of the linear operator $T(\boldsymbol{x}) = A\boldsymbol{x}$ is (Axler theorem 7.51)

$$T(\boldsymbol{v}) = s_1 \langle \boldsymbol{v}, \boldsymbol{e}_1 \rangle \boldsymbol{f}_1 + \cdots + s_n \langle \boldsymbol{v}, \boldsymbol{e}_n \rangle \boldsymbol{f}_n$$

where $s_k \geq 0$, and $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ and $\boldsymbol{f}_1, , \dots \boldsymbol{f}_n$ are orthonormal bases. Let $\boldsymbol{v} \in \mathbb{C}^n$. Writing the inner product in $\mathbb{C}^n$ as $\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \boldsymbol{u}^*\boldsymbol{v}$, we have for each $k = 1, \dots, n$,

$$s_k \langle \boldsymbol{v}, \boldsymbol{e}_k \rangle \boldsymbol{f}_k = s_k (\boldsymbol{e}_k^* \boldsymbol{v}) \boldsymbol{f}_k = s_k \boldsymbol{f}_k \boldsymbol{e}_k^* \boldsymbol{v}$$

thus

$$A\boldsymbol{v} = \sum_{k=1}^{n} s_k \boldsymbol{f}_k \boldsymbol{e}_k^* \boldsymbol{v} = \left(\sum_{k=1}^{n} s_k \boldsymbol{f}_k \boldsymbol{e}_k^* \right) \boldsymbol{v},$$

and since $\boldsymbol{v}$ was arbitrary,

$$A = \sum_{k=1}^{n} s_k \boldsymbol{f}_k \boldsymbol{e}_k^*.$$

Problem 6. Let A be a real 3×3 symmetric matrix, whose eigenvalues are λ_1, λ_2 and λ_3. Prove the following:

(a) If the trace of A, $\mathrm{tr}A$, is not an eigenvalue of A, then $(\lambda_1+\lambda_2)(\lambda_2+\lambda_3)(\lambda_1+\lambda_3) \neq 0$

(b) If $(\lambda_1 + \lambda_2)(\lambda_2 + \lambda_3)(\lambda_1 + \lambda_3) \neq 0$, then the map $T : S \to S$ is an isomorphism, where S is the space of 3×3 real skew-symmetric matrices (if $W^T = -W$, then W is called skew-symmetric) and $T(W) = AW + WA$.

Solution

(a) Since A is symmetric, A must be similar to a diagonal matrix $\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$. We also have $\mathrm{tr}A = \lambda_1 + \lambda_2 + \lambda_3$. The characteristic polynomial of A is $p(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3)$. If $\mathrm{tr}(A)$ is not an eigenvalue of A, $p(\mathrm{tr}A) \neq 0$, which means $(\lambda_1 + \lambda_2)(\lambda_2 + \lambda_3)(\lambda_1 + \lambda_3) \neq 0$.

(b) Since T is an operator on S, it suffices to show that T is injective. Let $T(W) = 0$ where $W^T = -W$. Then $AW = -WA$. So now we want to show $W = 0$.

Let $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ be the ordered basis of $\mathbb{R}^3$ with respect to which the representation of A is $\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$. We will now show $W\boldsymbol{a}_i = 0$, for $i = 1, 2, 3$, and hence $W = 0$.

Using the standard inner product in $\mathbb{R}^3$, we obtain, for $i = 1, 2, 3$,

$$\langle W\boldsymbol{a}_i, \boldsymbol{a}_i\rangle = \langle \boldsymbol{a}_i, W^T\boldsymbol{a}_i\rangle = -\langle \boldsymbol{a}_i, W\boldsymbol{a}_i\rangle = -\langle W\boldsymbol{a}_i, \boldsymbol{a}_i\rangle$$

So $W\boldsymbol{a}_i \perp \boldsymbol{a}_i$.

For $i \neq j$,

$$\langle AW\boldsymbol{a}_i, \boldsymbol{a}_j\rangle = -\langle WA\boldsymbol{a}_i, \boldsymbol{a}_j\rangle = \langle A\boldsymbol{a}_i, W\boldsymbol{a}_j\rangle = \lambda_i\langle \boldsymbol{a}_i, W\boldsymbol{a}_j\rangle$$

and

$$\begin{aligned}\langle AW\boldsymbol{a}_i, \boldsymbol{a}_j\rangle &= \langle W\boldsymbol{a}_i, A\boldsymbol{a}_j\rangle \\ &= \langle W\boldsymbol{a}_i, \lambda_j\boldsymbol{a}_j\rangle = \lambda_j\langle W\boldsymbol{a}_i, \boldsymbol{a}_j\rangle = -\lambda_j\langle \boldsymbol{a}_i, W\boldsymbol{a}_j\rangle\end{aligned}$$

So we have $\lambda_i\langle \boldsymbol{a}_i, W\boldsymbol{a}_j\rangle = -\lambda_j\langle \boldsymbol{a}_i, W\boldsymbol{a}_j\rangle$, i.e. $(\lambda_i + \lambda_j)\langle \boldsymbol{a}_i, W\boldsymbol{a}_j\rangle = 0$. From (b) we know $\lambda_i + \lambda_j \neq 0$, so we conclude $W\boldsymbol{a}_i \perp \boldsymbol{a}_j$, for $i \neq j$. So $W = 0$.

Problem 7. **<u>Note:</u> A similar problem was given in the UC Berkeley prelim, Spring 2021 as Problem 6A.**

1. Let A be an n-by-n real matrix such that all entries not on the diagonal are positive, and the sum of the entries in each row is negative.

   Calling C_j the j-th column of A, we perform the sequence of column operations:

$$\begin{array}{rcl} C_2 & \leftarrow & C_2 - \dfrac{a_{12}}{a_{11}} C_1 \\ C_3 & \leftarrow & C_3 - \dfrac{a_{13}}{a_{11}} C_1 \\ \vdots & & \vdots \\ C_n & \leftarrow & C_n - \dfrac{a_{1n}}{a_{11}} C_1. \end{array}$$

   So starting from A, we compute A' with this sequence of operation.

   Note that $a_{11} \neq 0$, so dividing by a_{11} makes sense. Also note that this sequence of column operations introduces 0 in the first row of A for all non diagonal elements. So we have

$$A' = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22}^{(1)} & a_{23}^{(1)} & \dots & a_{2n}^{(1)} \\ a_{31} & a_{32}^{(1)} & a_{33}^{(1)} & \dots & a_{3n}^{(1)} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2}^{(1)} & a_{n3}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}$$

Where

$$\begin{array}{rcl} & & \text{second column} \\ a_{22}^{(1)} & = & a_{22} - \frac{a_{12}}{a_{11}} a_{21} \\ a_{32}^{(1)} & = & a_{32} - \frac{a_{12}}{a_{11}} a_{31} \\ \cdots & & \\ a_{n2}^{(1)} & = & a_{n2} - \frac{a_{12}}{a_{11}} a_{n1} \\ & & \text{third column} \\ a_{23}^{(1)} & = & a_{23} - \frac{a_{13}}{a_{11}} a_{21} \\ a_{33}^{(1)} & = & a_{33} - \frac{a_{13}}{a_{11}} a_{31} \\ \cdots & & \\ a_{n3}^{(1)} & = & a_{n3} - \frac{a_{13}}{a_{11}} a_{n1} \\ & & \text{and so on} \end{array}$$

In short for $2 \le i \le n$ and $2 \le j \le n$, we have

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{1j}}{a_{11}} a_{i1}.$$

Let us call $A^{(1)}$ the $(n-1)$-by-$(n-1)$ matrix

$$A^{(1)} = \begin{pmatrix} a_{22}^{(1)} & a_{23}^{(1)} & \dots & a_{2n}^{(1)} \\ a_{32}^{(1)} & a_{33}^{(1)} & \dots & a_{3n}^{(1)} \\ \vdots & \vdots & & \vdots \\ a_{n2}^{(1)} & a_{n3}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}$$

Prove that $A^{(1)}$ is an $(n-1)$-by-$(n-1)$ real matrix such that all entries not on the diagonal are positive, and the sum of the entries in each row is negative.

2. Let A be an n-by-n real matrix such that all entries not on the diagonal are positive, and the sum of the entries in each row is negative. Show that the determinant of A is nonzero.

(Note: you can prove the second by assuming the first.)

Solution We will do a proof by induction with induction hypothesis

induction hypothesis All n-by-n matrix A with the property "all entries not on the diagonal are positive, and the sum of the entries in each row is negative," have a nonzero determinant.

For $n = 1$, the induction hypothesis is true since if $A = (a_{11})$ and "the sum of the entries in each row is negative", then $a_{11} < 0$, and so $\det(A) = a_{11} < 0$ and $\det(A)$ is nonzero.

Now let us assume the induction hypothesis true at step $n-1$, and prove that it is true at step n.

Let A be an n-by-n real matrix such that all entries not on the diagonal are positive, and the sum of the entries in each row is negative. We want to prove that the determinant of A is nonzero.

Our induction hypothesis is that we will assume that the statement "for all $(n-1)$-by-$(n-1)$ real matrix B such that all entries not on the diagonal are positive, and the sum of the entries in each row is negative, then the determinant of B is nonzero" is true.

Let's use the following notations

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix}$$

Since "all entries not on the diagonal are positive, and the sum of the entries in each row is negative", we have $a_{11} < 0$ and

$$-a_{11} > a_{12} + a_{13} + \ldots + +a_{1n}.$$

In particular a_{11} is nonzero.

Calling C_j the j-th column of A. We perform the sequence of column operations:

$$\begin{array}{rcl} C_2 & \leftarrow & C_2 - \dfrac{a_{12}}{a_{11}} C_1 \\ C_3 & \leftarrow & C_3 - \dfrac{a_{13}}{a_{11}} C_1 \\ \vdots & & \vdots \\ C_n & \leftarrow & C_n - \dfrac{a_{1n}}{a_{11}} C_1. \end{array}$$

So starting from A, we compute A' with this sequence of operation.

Note that $a_{11} \neq 0$, so dividing by a_{11} makes sense. Also note that all these operations conserve the determinant. So $\det(A) = \det(A')$. Also note that this sequence of column operations introduces 0 in the first row of A for all non diagonal elements. So we have

$$A' = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22}^{(1)} & a_{23}^{(1)} & \dots & a_{2n}^{(1)} \\ a_{31} & a_{32}^{(1)} & a_{33}^{(1)} & \dots & a_{3n}^{(1)} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2}^{(1)} & a_{n3}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}$$

where

$$\begin{aligned} & & \text{second column} \\ a_{22}^{(1)} &= & a_{22} - \frac{a_{12}}{a_{11}} a_{21} \\ a_{32}^{(1)} &= & a_{32} - \frac{a_{12}}{a_{11}} a_{31} \\ \dots & & \\ a_{n2}^{(1)} &= & a_{n2} - \frac{a_{12}}{a_{11}} a_{n1} \\ & & \text{third column} \\ a_{23}^{(1)} &= & a_{23} - \frac{a_{13}}{a_{11}} a_{21} \\ a_{33}^{(1)} &= & a_{33} - \frac{a_{13}}{a_{11}} a_{31} \\ \dots & & \\ a_{n3}^{(1)} &= & a_{n3} - \frac{a_{13}}{a_{11}} a_{n1} \\ & & \text{and so on} \end{aligned}$$

In short for $2 \le i \le n$ and $2 \le j \le n$, we have

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{1j}}{a_{11}} a_{i1}.$$

Let us call $A^{(1)}$ the $(n-1)$-by-$(n-1)$ matrix

$$A^{(1)} = \begin{pmatrix} a_{22}^{(1)} & a_{23}^{(1)} & \dots & a_{2n}^{(1)} \\ a_{32}^{(1)} & a_{33}^{(1)} & \dots & a_{3n}^{(1)} \\ \vdots & \vdots & & \vdots \\ a_{n2}^{(1)} & a_{n3}^{(1)} & \dots & a_{nn}^{(1)} \end{pmatrix}$$

We will prove below that $A^{(1)}$ has the property "all entries not on the diagonal are positive, and the sum of the entries in each row is negative."

Let us assume for now that $A^{(1)}$ has the property "all entries not on the diagonal are positive, and the sum of the entries in each row is negative." Since $A^{(1)}$ is $(n-1)$-by-$(n-1)$ matrix with the property "all entries not on the diagonal are positive, and the sum of the entries in each row is negative.” By induction we know that the determinant of $A^{(1)}$ is nonzero.

Now we understand that (by expanding the determinant with the first row of A', and since $\det(A) = \det(A')$)

$$\det(A) = a_{11} \times \det(A^{(1)}).$$

Since a_{11} and $\det(A^{(1)})$ are nonzero, we conclude that

$$\text{the determinant of } A \text{ is nonzero.}$$

And that concludes our proof.

Now it remains to prove that $A^{(1)}$ has the property “all entries not on the diagonal are positive, and the sum of the entries in each row is negative.”

Firstly, let us prove that

$$\text{all off-diagonal elements of } A^{(1)} \text{ are positive.}$$

Indeed let i be such that $2 \leq i \leq n$ and j such that $2 \leq j \leq n$, with $i \neq j$. Then

1. a_{i1} is not a diagonal element of A, (since $i \neq 1$,) so $a_{i1} > 0$;
2. a_{1j} is not a diagonal element of A, (since $j \neq 1$,) so $a_{1j} > 0$;
3. a_{11} is a diagonal element of A, so $a_{11} < 0$;

so that

$$-\frac{a_{1j}}{a_{11}} a_{i1} > 0$$

so that

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{1j}}{a_{11}} a_{i1} > a_{ij}.$$

Now, a_{ij} is not a diagonal element of A, (since $i \neq j$,) so $a_{ij} > 0$. Since $a_{ij} > 0$, and $a_{ij}^{(1)} > a_{ij}$, we have

$$a_{ij}^{(1)} > 0.$$

Secondly, let us prove that

$$\text{the sum of the entries in each row of } A^{(1)} \text{ is negative.}$$

Indeed let i be such that $2 \leq i \leq n$. Then since the sum of the entries in each row of A is negative, for the first row we get

$$a_{11} + a_{12} + a_{13} + \ldots + a_{1n} < 0$$

Multiplying by $-\frac{a_{i1}}{a_{11}}$, (which is a positive number,) we get

$$-\frac{a_{i1}}{a_{11}}a_{11} - \frac{a_{i1}}{a_{11}}a_{12} - \frac{a_{i1}}{a_{11}}a_{13} + \ldots - \frac{a_{i1}}{a_{11}}a_{1n} < 0$$

Since the sum of the entries in each row of A is negative, for the i-th row we get

$$a_{i1} + a_{i2} + a_{i3} + \ldots + a_{in} < 0$$

Adding the two last equations together gives

$$(a_{i1} - \frac{a_{i1}}{a_{11}}a_{11}) + (a_{i2} - \frac{a_{i1}}{a_{11}}a_{12}) + (a_{i3} - \frac{a_{i1}}{a_{11}}a_{13}) + \ldots + (a_{in} - \frac{a_{i1}}{a_{11}}a_{1n}) < 0$$

which gives

$$a_{i2}^{(1)} + a_{i3}^{(1)} + \ldots + a_{in}^{(1)} < 0$$

We see that the sum of the entries in row $i-1$ of $A^{(1)}$ is negative. Since i was arbitrary between 2 and n, this proves the result.

Comment:

1. It is clear that we can prove that the sign of the determinant alternate. If n is odd, A is negative. If n is even, A is positive. The question only asked to prove nonzero, so we just proved nonzero.
2. The proof is somewhat related to the fact that LU with partial pivoting does not pivot for diagonally dominant matrices.

Problem 8.

Note: A similar problem was given in the UC Berkeley prelim, Fall 2020 as Problem 6B.

Let n be an integer.

We consider the inner product space, $\mathcal{P}_n$, of the polynomials of degree at most n with the inner product

$$< P, Q >= \int_0^1 P(t)Q(t)dt.$$

(Where $P(t)$ and $Q(t)$ are two polynomials of degree at most n.)

We consider $\{P_0, P_1, P_2, \ldots, P_n\}$ an orthonormal basis of $\mathcal{P}_n$, (orthonormal with respect to the aforementioned inner product), and so we note that we have

$$\int_0^1 P_i(t)P_j(t)dt = \delta_{ij}, \quad 0 \le i, j \le n.$$

Let H be the $(n+1) \times (n+1)$ Hilbert matrix. H is defined such that its elements are

$$h_{ij} = \int_0^1 t^i t^j dt, \quad 0 \le i, j \le n.$$

(Note that indexing starts at 0.)

We define the $(n+1) \times (n+1)$ matrix P made of the coefficients of the orthonormal basis $\{P_0, P_1, P_2, \ldots, P_n\}$ in the basis $\{1, t, t^2, \ldots, t^n\}$. So P is such that p_{ij} is the j-th coefficient of $P_i(t)$ so that

$$P_i(t) = \sum_{j=0}^{n} p_{ij} t^j, \quad 0 \le i \le n.$$

(Note that, with our notation, P is a matrix, and P_i is a polynomial, and p_{ij} is a coefficient of P_i and of P.)

Show that $H^{-1} = P^T P$.

Solution Let i be such that $0 \le i \le n$ and j such that $0 \le j \le n$; we have

$$\begin{aligned}
\text{entry } (i,j) \text{ of } I &= \delta_{i,j} \\
&= \int_0^1 P_i(t)P_j(t)dt \\
&= \int_0^1 (\sum_{k=0}^{n} p_{ik}t^k)(\sum_{l=0}^{n} p_{jl}t^l)dt \\
&= \sum_{k=0}^{n} \left(p_{ik} \left(\sum_{l=0}^{n} p_{jl} \left(\int_0^1 t^k t^l dt \right) \right) \right) \\
&= \sum_{k=0}^{n} \left(p_{ik} \left(\sum_{l=0}^{n} p_{jl} H_{kl} \right) \right) \\
&= \sum_{k=0}^{n} \left(p_{ik} \left(\text{entry } (k,j) \text{ of } HP^T \right) \right) \\
&= \text{entry } (i,j) \text{ of } PHP^T
\end{aligned}$$

So that

$$I = PHP^T$$

Since I is invertible, so are P, H and P^T and applying P^{-1} and $(P^{-1})^T = P^{-T}$ gives

$$P^{-1}P^{-T} = H.$$

Inverting both sides gives

$$H^{-1} = P^T P.$$

Couple comments:

1. The setting of this problem is with $\mathcal{P}_n$ and the inner product the $\langle P, Q\rangle = \int_0^1 P(t)Q(t)dt$. The Hilbert matrix happens to be here because one should note that the Hilbert matrix is the matrix of the inner products of the (non-orthonormal) basis $\{1, t, t^2, \dots, t^n\}$. And our problem is set in the (non-orthonormal) basis $\{1, t, t^2, \dots, t^n\}$.

   For any polynomials, say $P(t)$ and $Q(t)$, if you write them as vectors, say x and y, in the basis $\{1, t, t^2, \dots, t^n\}$, and you want to compute the inner product $\langle P, Q\rangle = \int_0^1 P(t)Q(t)dt$, you can do
   $$\langle P, Q\rangle = x^T H^{-1} y.$$
   So the problem is essentially asking to redemonstrate this property.

2. If one take, for $i = 0$ to n, $P_i(t)$ to be the Legendre polynomial of degree i. The Legendre polynomials are defined as follows. We consider the inner product space, $\mathcal{P}_n$, of the polynomial of degree at most n with the inner product
   $$< P, Q >= \int_0^1 P(t)Q(t)dt.$$
   (Where $P(t)$ and $Q(t)$ are two polynomials of degree at most n.) Then we perform Gram-Schmidt algorithm on the basis $\{1, x, x^2, \dots, x^n\}$ to obtain the Legendre polynomials $\{P_0, P_1, P_2, \dots, P_n\}$. The Legendre polynomial form an orthonormal basis of $\mathcal{P}_n$. But we also have the fact that
   $$\begin{aligned} &\text{Span}(1) = \text{Span}(P_0)\\ &\text{Span}(1, x) = \text{Span}(P_0, P_1)\\ &\text{Span}(1, x, x^2) = \text{Span}(P_0, P_1, P_2) \qquad (1)\\ &\vdots \end{aligned}$$
   If then, we denote p_{ij} the j-th coefficient of the Legendre polynomial of degree i so that
   $$P_i(t) = \sum_{j=0}^{i} p_{ij} t^j, \quad 0 \le i \le n.$$

We note that, thanks to Equation (1), P is upper triangular. (And this is why the sum in the equation above stops early at i and does not have to go all the way n.) Therefore, in this problem, proving that $H^{-1} = P^T P$ is proving that "the Cholesky factor of the inverse of the Hilbert matrix is the matrix with coefficients same as the coefficients of the Legendre polynomials."