MIDTERM EXAM: TAKE-HOME PART
MATH 4070, FALL 2025

This take-home exam is open book, open notes, and closed everything else. You may refer to any of the following:

- the lecture notes (your own or the ones I upload to Canvas),
- Prof. Keinert's notes,
- the optional textbook (Helene Shapiro, *Linear algebra and matrices*),
- your introductory linear algebra textbook (most likely Lay or Strang),
- your own homework for this class, including the grader's comments, and
- MATLAB's internal documentation.

You may not refer to anything else. In particular:

- you may not use AI,
- you may not use the internet in any way, and
- you may not talk to anybody else about these problems.

If you have any questions about the rules, **ask!**

Throughout this exam, $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ and $\|\cdot\|$ is the norm given by the (standard) inner product.

At least one of the following problems will appear on the in-class midterm. For any such problems, only your score on the in-class midterm will count.

Problems to type in TEX

1. Fix integers $m, n, p \geq 1$ and a matrix $A \in \mathbb{F}^{n\times p}$, and consider the function $L\colon \mathbb{F}^{m\times n} \to \mathbb{F}^{m\times p}$ given by $L(X) = XA$ for $X \in \mathbb{F}^{m\times n}$.
   (a) Verify that L is linear.

   It suffices to show that L is linear if it holds both properties for any matrices $X_1, X_2 \in \mathbb{F}^{m\times n}$ and any scalar $c \in \mathbb{F}$: $L(X_1+X_2) = L(X_1)+L(X_2)$ and $L(cX_1) = cL(X_1)$. First, $L(X_1 + X_2) = (X_1 + X_2)A = X_1A+X_2A = L(X_1)+L(X_2)$. Thus, the first property holds. Second, $L(cX_1) = (cX_1)A = c(X_1A) = cL(X_1)$. Thus, both properties hold $\implies$ L is linear.

   (b) Using the Frobenius inner product, verify that the adjoint transformation $L^*\colon \mathbb{F}^{m\times p} \to \mathbb{F}^{m\times n}$ is given by $L^*(Y) = YA^*$ for $Y \in \mathbb{F}^{m\times p}$.

   By the definition of an adjoint transformation, $\langle L(X), Y\rangle_F = \langle X, L^*(Y)\rangle_F$. Then, we simplify the left hand side and use the definition of Frobenius inner product, $\langle L(X), Y\rangle_F = \langle XA, Y\rangle_F = tr((XA)^*Y) = tr(A^*X^*Y)$. Then, by the cyclic property of traces, $tr(A^*X^*Y) = tr(X^*YA^*)$. Going back to our right hand side, by the definition of Frobenius inner product, $\langle X, L^*(Y)\rangle_F = tr(X^*L^*(Y))$. So, we must have $tr(X^*YA^*) = tr(X^*L^*(Y)) \implies tr(X^*YA^*)-tr(X^*L^*(Y)) = 0 \implies tr(X^*(YA^*-L^*(Y))) = 0$. Let $B = YA^*-L^*(Y)$, note that B does exist since YA^*

is m by n and so is $L^*(Y)$. We therefore want $tr(X^*B) = 0$ for all $X \in \mathbb{F}^{m\times n}$. This can easily be observed as only being true when B is the zero matrix (since $X = B$ is a possibility. So, $B = 0 \implies YA^* - L^*(Y) = 0 \implies L^*(Y) = YA^*$. Thus, the statement is true.

(c) Show that if $p = n$ and $\lambda \in \mathbb{F}$ is an eigenvalue of A, then exists a nonzero matrix $X \in \mathbb{F}^{m\times n}$ for which $L(X) = \lambda X$.

Since $p = n$, $A \in \mathbb{F}^{n\times n}$. We know λ is an eigenvalue of A. So, by definition of an eigenvalue, $Av = \lambda v$ for some non-zero eigenvector $v \in \mathbb{F}^n$. Since A is square, λ is an eigenvalue for A^T, such that $A^T u = \lambda u$ for some non-zero eigenvector $u \in \mathbb{F}^n$. Taking the transpose, $(A^T u)^T = (\lambda u)^T \implies u^T A = \lambda u^T$. Let $x = u^T$ such that $xA = \lambda x$. Let the matrix X have it's first row be the non-zero row vector x, with the other rows being 0 vectors. So, $L(X) = XA$, where the first row of XA is xA and the rest are the 0 vector. We know $xA = \lambda x$, so XA has its first row as λx and the rest are the 0 vector. Since λ is a scalar, taking it out results in the first row being x, which is just X. Thus, $L(X) = \lambda X$. So, the statement is true.

(d) Conversely, show that if $p = n$ and for some $\lambda \in \mathbb{F}$ there exists a nonzero matrix $X \in \mathbb{F}^{m\times n}$ for which $L(X) = \lambda X$, then λ is an eigenvalue of A.

Since $p = n$, $A \in \mathbb{F}^{n\times n}$. It suffices to show $det(A - \lambda I) = 0$. We know $L(X) = \lambda X \implies XA = \lambda X \implies X(A - \lambda I) = 0$. We know X is non-zero, so there exists at least one row x that is not a zero vector in X. This means for any non-zero row vector $x \in X$, $x(A - \lambda I) = 0$. Thus, there exists a vector x in the null space of $(A - \lambda I)$. This means $(A - \lambda I)$ is non-invertible. Since $(A - \lambda I)$ is non-invertible and square, it's determinant must be 0. Thus, the statement is true.

2. Let V and W be finite-dimensional inner product spaces over $\mathbb{F}$ (not necessarily $\mathbb{F}^n$), and let $L\colon V \to W$ be a linear transformation. Given a desired output $b \in W$, consider the problem of finding an input $\hat{x} \in V$ that minimizes $\|L(\hat{x}) - b\|^2$. Prove that the set of such *least-squares solutions* to the system $L(x) = b$ coincides with the set of all $\hat{x} \in V$ for which $L^*(L(\hat{x})) = L^*(b)$.

Let $S_1 = \{\hat{x} \in V : ||L(\hat{x}) - b||^2 \leq ||L(x) - b||^2 \text{ for all } x \in V\}$ be the set of least-squares solutions for the system $L(x) = b$ and $S_2 = \{\hat{x} \in V : L^*(L(\hat{x})) = L^*(b)\}$ be the set of solutions to norm equation. For two sets to be equal, they must be subsets of each other. So, it suffices to show $S_1 \subseteq S_2$ and $S_2 \subseteq S_1$.

First, we show $S_1 \subseteq S_2$. Let $\hat{x} \in S_1$. We know $\hat{x}$ minimizes $L(x) - b$. So, by definition, the vector $L(\hat{x})$ is a point in $\mathrm{Im}(L) \subseteq W$ that is closest to the vector $b \in W$. This means, by an orthogonal projection, $b - L(\hat{x})$ must be orthogonal to $\mathrm{Im}(L)$. Thus, we can now say $\langle L(v), b - L(\hat{x})\rangle = 0$ for any $v \in V$. This can be rearranged using adjoints, $\langle L(v), b - L(\hat{x})\rangle = 0 \implies \langle v, L^*(b - L(\hat{x})) = 0$. This applies to all vectors in V, and the only vector orthogonal to every vector in V is the zero vector. So, we want

$L^*(b - L(\hat{x})) = 0 \implies L^*(b) - L^*(L(\hat{x})) = 0 \implies L^*(b) = L^*(L(\hat{x}))$. This is equivalent to the requirements for S_2, thus $\hat{x} \in S_2$. This implies then that $S_1 \subseteq S_2$.

Then, we show $S_2 \subseteq S_1$. Let $\hat{x} \in S_2$, so we know $L^*(L(\hat{x})) = L^*(b)$. This can be rearranged, $L^*(L(\hat{x})) = L^*(b) \implies L^*(L(\hat{x})) - L^*(b) = 0 \implies L^*(L(\hat{x}) - b) = 0$. We can say for any $u \in \text{Im}(L)$ that $u = L(v)$ for some $v \in V$. Then, $\langle u, L(\hat{x}) - b\rangle = \langle L(v), L(\hat{x}) - b\rangle = \langle v, L^*(L(\hat{x}) - b)\rangle = \langle v, 0\rangle = 0$. Therefore, $L(\hat{x}) - b$ is orthogonal to $\text{Im}(L)$. We now want to show $||L(\hat{x}) - b||^2 \leq ||L(x) - b||^2$ for any vector $x \in V$. Rewriting the equation, $||L(\hat{x}) - b||^2 \leq ||L(x) - b||^2 \implies ||L(\hat{x}) - b||^2 \leq ||(L(x) - L(\hat{x})) + (L(\hat{x}) - b)||^2$. Since L is linear, $L(x) - L(\hat{x}) = L(x - \hat{x})$, which must be in $\text{Im}(L)$ since $x - \hat{x} \in V$. We also know $L(\hat{x}) - b$ is orthogonal to $\text{Im}(L)$. We therefore can say $L(x - \hat{x})$ is orthogonal to $L(\hat{x}) - b$. Then, by the Pythagorean Theorem (since they are orthogonal), $||L(\hat{x}) - b||^2 \leq ||(L(x - \hat{x}) + (L(\hat{x}) - b)||^2 \implies ||L(\hat{x}) - b||^2 \leq ||(L(x - \hat{x})||^2 + ||L(\hat{x}) - b||^2$. Which, by basic observation, must hold since the norm must be non-negative. Thus, $\hat{x}$ minimizes $||L(\hat{x}) - b||^2$ and therefore is in S_1. This proves, $S_2 \subseteq S_1$.

Since $S_1 \subseteq S_2$ and $S_2 \subseteq S_1$, we can conclude they are equal. Thus, they coincide with each other.

3. Given desired input vectors $a_1, \dots, a_N \in \mathbb{F}^n$ and corresponding outputs $b_1, \dots, b_N \in \mathbb{F}^m$, consider the problem of finding a matrix $\hat{X} \in \mathbb{F}^{m \times n}$ that minimizes $\sum_{j=1}^{N} \|\hat{X}a_j - b_j\|^2$. Assume that $\text{span}\{a_1, \dots, a_N\} = \mathbb{F}^n$, and denote

$$A = \begin{bmatrix} a_1 & \cdots & a_N \end{bmatrix} \in \mathbb{F}^{n \times N}, \qquad B = \begin{bmatrix} b_1 & \cdots & b_N \end{bmatrix} \in \mathbb{F}^{m \times N},$$

and

$$B^* = \begin{bmatrix} r_1 & \cdots & r_m \end{bmatrix} \in \mathbb{F}^{N \times m}.$$

Show that the problem above has a unique solution $\hat{X}$, and the columns of

$$\hat{X}^* = \begin{bmatrix} \hat{x}_1 & \cdots & \hat{x}_m \end{bmatrix} \in \mathbb{F}^{n \times m}$$

satisfy $AA^*\hat{x}_i = Ar_i$ for every $i = 1, \dots, m$.

Hint: Apply Problem 2 with $V = \mathbb{F}^{m \times n}$, $W = \mathbb{F}^{m \times N}$, and $L(X) = XA$.

Let $V = \mathbb{F}^{m \times n}$, $W = \mathbb{F}^{m \times N}$. Let $X \in V$. The columns of matrix product $XA \in W$ are $Xa_1, \dots, Xa_N$. Similarly, for $B \in W$ are $b_1, \dots, b_N$. We then can say $XA - B = [Xa_1 - b_1 \dots Xa_N - b_N]$. Taking the Frobenius norm of this, $||XA - B||_F^2 = \sum_{j=1}^{N} |Xa_j - b_j|^2$. Thus, this is equivalent to finding some matrix $\hat{X} \in V$ that minimizes $||L(\hat{X}) - B||_F^2$, where $L : V \to W$ is the linear transformation $L(X) = XA$. From problem 2, the least-squares solution $\hat{X}$ must satisfy the equation $L^*(L(\hat{X})) = L^*(B)$. First, lets figure out the adjoint $L^* : W \to V$. For any $X \in V$, $Y \in W$, the Frobenius inner product is $\langle L(X), Y\rangle_F = \langle XA, Y\rangle_F$. Then, by the definition of Frobenius inner product, $\langle XA, Y\rangle_F = tr(Y^*XA)$. Then, by trace cycles, $tr(Y^*XA) = tr(AY^*X)$. Compare this with the other side of the adjoint definition, $\langle X, L^*(Y)\rangle_F = tr((L^*(Y))^*X)$. So, $tr(AY^*X) =$

$tr((L^*(Y))^*X)$. Thus, we can say $(L^*(Y))^* = AY^* \implies L^*(Y) = YA^*$. Applying this to $L^*(L(\hat{X})) = L^*(B)$, we get $\hat{X}AA^* = BA^*$.

Next, lets prove the uniqueness of $\hat{X}$ for $\hat{X}AA^* = BA^*$. We need AA^* to be invertible. We know it is a square matrix and rank$(A) = n$ (since its columns span $\mathbb{F}^n$). By Prof. Keinert's notes (the theorem is unnamed), rank(AA^*) = rank$(A) = n$. Thus, because each column is linearly independent, AA^* is invertible. We can now try solving the equation by rearranging, $\hat{X}AA^* = BA^* \implies \hat{X} = BA^*(AA^*)^{-1}$. Therefore, there is a unique solution $\hat{X}$.

Finally, we have show the columns of $\hat{X}^*$ satisfy $AA^*\hat{x}_i = Ar_i$. Rearranging the equation we found earlier, $(\hat{X}AA^*)^* = (BA^*)^* \implies AA^*\hat{X}^* = AB^*$. We know, $\hat{X}^* = \begin{bmatrix} \hat{x}_1 & \cdots & \hat{x}_m \end{bmatrix} \in \mathbb{F}^{n\times m}$ and $B^* = \begin{bmatrix} r_1 & \cdots & r_m \end{bmatrix} \in \mathbb{F}^{N\times m}$. The ith column of $AA^*\hat{X}^*$ is easily observed as AA^* multiplied by the ith column of $\hat{X}^*$, or $AA^*\hat{x}_i^*$. The ith column of AB^* is easily observed as A multiplied as the ith column of B^*, or Ar_i. Putting this together, $AA^*\hat{X}^* = AB^* \implies AA^*\hat{x}_i = Ar_i$. This holds for every $i = 1, \ldots, m$. Thus satisfying our condition.

Both conditions hold, finishing our proof.

Problems to do in MATLAB

4. The effects of a linear transformation $T\colon \mathbb{R}^{150} \to \mathbb{R}^{100}$ were observed under noisy conditions by recording inputs $a_1, \cdots, a_{2000} \in \mathbb{R}^{150}$ and their observed outputs $b_1, \cdots, b_{2000} \in \mathbb{R}^{100}$. These were stored as columns of the matrices $A \in \mathbb{R}^{150\times 2000}$ and $B \in \mathbb{R}^{100\times 2000}$, respectively, in the file `Midterm_Data_1.mat`.
   (a) Estimate a matrix $X \in \mathbb{R}^{100\times 150}$ for T using the method of Problem 3. Specifically, employ a Cholesky decomposition to solve the equations $AA^*\hat{x}_i = Ar_i$.
   (b) Improve your estimate for the matrix of T as follows. First, use the singular values of X to estimate the dimension r of the image of T. Then, find the nearest matrix Y to X with rank r.
   (c) The file `Midterm_Data_1.mat` also contains a desired output vector $c \in \mathbb{R}^{100}$. Use the "approximate QR method" with Y to predict an input $x \in \mathbb{R}^{150}$ of small norm for which $T(x) \approx c$. Show all the details; do not just use the "\" command.

5. The file `Midterm_Data_2.mat` contains a matrix $B \in \mathbb{R}^{10\times 10}$. Use MATLAB to answer the following question: Does the function $Q\colon \mathbb{R}^{10} \to \mathbb{R}$ given by
$$Q(x) = x^\top Bx$$
ever take negative values on $\mathbb{R}^{10}$? If yes, give an explicit x for which $Q(x)$ is negative. If no, explain in the comments how you can tell.