University of Colorado Denver
Department of Mathematical and Statistical Sciences
Applied Linear Algebra Ph.D. Preliminary Exam
January 23, 2015

Name: _____

Exam Rules:

- This exam lasts 4 hours and consists of 6 problems worth 20 points each.

- Each problem will be graded, and your final score will count out of 120 points.

- You are not allowed to use your books or any other auxiliary material on this exam.

- Start each problem on a separate sheet of paper, write only on one side, and label all of your pages in consecutive order (*e.g.*, use 1-1, 1-2, 1-3, ..., 2-1, 2-2, 2-3, ...).

- Read all problems carefully, and write your solutions legibly using a dark pencil or pen "essay-style" using full sentences and correct mathematical notation.

- Justify all your solutions: cite theorems you use, provide counterexamples for disproof, give clear but concise explanations, and show calculations for numerical problems.

- If you are asked to prove a theorem, you may not merely quote or rephrase that theorem as your solution; instead, you must produce an independent proof.

- If you feel that any problem or any part of a problem is ambiguous or may have been stated incorrectly, please indicate your interpretation of that problem as part of your solution. Your interpretation should be such that the problem is not trivial.

- Please ask the proctor if you have any other questions.

1. _____   4. _____

2. _____   5. _____

3. _____   6. _____

Total _____

DO NOT TURN THE PAGE UNTIL TOLD TO DO SO.

Applied Linear Algebra Preliminary Exam Committee:
Alexander Engau (Chair), Julien Langou, Anatolii Puhalskii.

Problem 1

Find a basis for the intersection of the subspace of $\mathbb{R}^4$ spanned by $(1,1,0,0)$, $(0,1,1,0)$, $(0,0,1,1)$ and the subspace spanned by $(1,0,t,0)$, $(0,1,0,t)$, where t is given. [20 points]

Solution: Let $S_1 = \text{Span} \left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right)$ and $S_2 = \text{Span} \left(\begin{pmatrix} 1 \\ 0 \\ t \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ t \end{pmatrix} \right)$.

Method 1: The two subspaces are given in parametric form. We may convert to Cartesian form and then solve for the intersection. Hence, we may begin to find a basis for $S_1^\perp$, *i.e.*, for the nullspace of

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

This matrix is in row echelon form with x_1, x_2, and x_3 as leading variables and with x_4 as a free variable. Hence, a possible basis for its nullspace is given by $\{(-1, 1, -1, 1)\}$, yielding the following Cartesian equation for S_1:

$$-x_1 + x_2 - x_3 + x_4 = 0.$$

Next, we need to find a basis for $S_2^\perp$, *i.e.*, for the nullspace of

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

This matrix is in row echelon form with x_1 and x_2 as leading variables and with x_3 and x_4 as free variables. Hence, a possible basis for its nullspace is given by $\{(-t, 0, 1, 0), (0, -t, 0, 1)\}$, yielding the following Cartesian equations for S_2:

$$\begin{aligned} -tx_1 \qquad\quad + x_3 \qquad\quad &= 0; \\ -tx_2 \qquad\quad + x_4 &= 0. \end{aligned}$$

It follows that the intersection of S_1 and S_2 can be represented in Cartesian form as

$$\begin{aligned} -x_1 + \quad x_2 - \quad x_3 + \quad x_4 &= 0; \\ -tx_1 \qquad\qquad + \quad x_3 \qquad\quad &= 0; \\ -tx_2 \qquad\qquad + \quad x_4 &= 0. \end{aligned}$$

Next, to find a basis, we convert back to matrix form and apply an elementary row operation:

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

We continue with two cases, based on whether $t \neq -1$ or $t = -1$.

1. Case 1: If $t \neq -1$, then $1 + t \neq 0$ and we can divide the first row by $1 + t$ to get
$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ -t & 0 & 1 & 0 \\ 0 & -t & 0 & 1 \end{pmatrix}.$$

It follows that x_2, x_3 and x_4 can be taken as leading variables with x_1 as a free variable, so that a basis for its nullspace is, for example: $\{(1, 1, t, t)\}$. This is a basis for $S_1 \cap S_2$, for the case that $t \neq -1$.

2. Case 2: If $t = -1$, then we get
$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

We see that x_1 and x_2 can be taken as leading variables with x_3 and x_4 as free variables, so that a basis for its nullspace is, for example: $\{(1, 0, -1, 0), (0, 1, 0, -1)\}$. This is a basis for $S_1 \cap S_2$, for the case that $t = -1$.

Method 2: We look for linearly dependent vectors. By row reduction,

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & t & 0 \\ 0 & 0 & 1 & 0 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & t & 0 \\ 0 & 0 & 1 & 0 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & t+1 & -1 \\ 0 & 0 & 1 & 0 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 \\ 0 & 0 & 1 & t+1 & -1 \\ 0 & 0 & 0 & -t-1 & t+1 \end{bmatrix}.$$

Since row reduction preserves linear dependence relations among columns, it follows that if $t = -1$, then the last two vectors of the original set are linear combinations of the first three. Hence, the intersection of the subspaces is the subspace spanned by the last two vectors, which are linearly independent, so a basis is $\{(1, 0, -1, 0), (0, 1, 0, -1)\}$.

If $t \neq -1$, then the fourth and the fifth vectors are not in the span of the first three, but each of them is in the span of the other four, of course. Hence, the intersection is a one-dimensional subspace. The sum of vectors four and five belongs to either subspace, for example:

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & t \\ 0 & 0 & 1 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & t \\ 0 & 0 & 1 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t \\ 0 & 0 & 1 & t \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence, a basis is $\{(1, 1, t, t)\}$.

Method 3: An element $x \in \mathbb{R}^4$ belongs to the intersection $S_1 \cap S_2$ if and only if

$$x = \sigma_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \sigma_2 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \sigma_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \sigma_1 \\ \sigma_1 + \sigma_2 \\ \sigma_2 + \sigma_3 \\ \sigma_3 \end{pmatrix} = \tau_1 \begin{pmatrix} 1 \\ 0 \\ t \\ 0 \end{pmatrix} + \tau_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ t \end{pmatrix} = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_1 t \\ \tau_2 t \end{pmatrix}$$

for scalars σ_1, σ_2, σ_3, and τ_1, τ_2. It follows that $\sigma_1 = \tau_1$, $\sigma_3 = \tau_2 t$, $\sigma_2 = \tau_2 - \sigma_1 = \tau_2 - \tau_1$, and

$$\sigma_3 = \tau_1 t - \sigma_2 = \tau_1 t - (\tau_2 - \tau_1) = \tau_1(t + 1) - \tau_2,$$

so that $\sigma_3 = \tau_2 t = \tau_1(t + 1) - \tau_2$ and thus $(\tau_1 - \tau_2)(t + 1) = 0$. We continue with two cases.

1. If $t = -1$, then τ_1 and τ_2 can be set arbitrarily, so that $S_1 \cap S_2 = S_2$ with basis vectors $(1, 0, -1, 0)$, $(0, 1, 0, -1)$ (the same basis as given for S_2 with $t = -1$).

2. If $t \neq -1$, then τ_1 and τ_2 must be set equal, so that $S_1 \cap S_2$ is spanned by $(1, 1, t, t)$.

Problem 2

Let A be a real $m \times n$ matrix and b be a real m-dimensional vector. Prove the following.

(a) If the equation $Ax = b$ is consistent, then there exists a unique vector $p \in \text{Row } A$ such that $Ap = b$. [10 points]

(b) The equation $Ax = b$ is consistent if and only if the vector b is orthogonal to every solution of $A^T y = 0$. [10 points]

Solution: A result that one may use is that the orthogonal complement of the null space is the row space, *i.e.*,
$$(\text{Nul } A)^\perp = \text{Row } A. \tag{*}$$
A direct consequence is that $\text{Nul } A \oplus \text{Row } A = \mathbb{R}^n$.

(a) Assume that $Ax = b$ is consistent. It follows that there exists x in $\mathbb{R}^n$ such that $Ax = b$. Since $\mathbb{R}^n = \text{Row } A \oplus \text{Nul } A$, there exists $p \in \text{Row } A$ and $y \in \text{Nul } A$ such that $x = p + y$, and thus $Ap = Ax = b$. This proves the existence of $p \in \text{Row } A$ such that $Ap = b$.

To prove that p is unique, let $p_1 \in \text{Row } A$ and $p_2 \in \text{Row } A$ such that $Ap_1 = Ap_2 = b$. It follows that $p_1 - p_2 \in \text{Row } A \cap \text{Nul } A$, which is the zero subspace, and thus $p_1 = p_2$.

(b) Since (1) $b \in \text{Col } A$ if and only if $Ax = b$ is consistent, and (2) $b \in (\text{Nul } A^T)^\perp$ if and only if b is orthogonal to every solution of $A^T y = 0$, the question amounts to prove that
$$\text{Col } A = (\text{Nul } A^T)^\perp.$$

This latter subspace equality can be seen by noting that (1) $\text{Col } A = \text{Row } A^T$ (clear) and (2) $\text{Row } A^T = (\text{Nul } A^T)^\perp$ (by using (**??**) on A^T).

Problem 3

Let A and B be $n \times n$ matrices. Prove or disprove each of the following. [5 points each]

(a) If A and B are diagonalizable, then so is $A + B$.

(b) If A and B are diagonalizable, then so is AB.

(c) If $A^2 = A$, then A is diagonalizable.

(d) If A^2 is diagonalizable, then so is A.

Solution:

(a) The statement is not true. A counterexample is

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix} \text{ with } A + B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

The matrix A is upper triangular, so that its eigenvalues are its diagonal entries 1 and 0. Since A has two distinct eigenvalues and is a 2×2 matrix, A is diagonalizable. Similarly, the matrix B is diagonal and thus diagonalizable. Because $A + B$ is a Jordan block of size 2 associated with eigenvalue 0, it follows that $A + B$ is not diagonalizable. Hence, we can see that A and B are both diagonalizable, while $A + B$ is not diagonalizable.

(b) The statement is not true. A counterexample is

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \text{ with } AB = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

The matrix A is upper triangular, so that its eigenvalues are its diagonal entries 1 and 0. Since A has two distinct eigenvalues and is a 2×2 matrix, A is diagonalizable. Similarly, the matrix B is diagonal and thus diagonalizable. Because $A + B$ is a Jordan block of size 2 associated with eigenvalue 0, it follows that $A + B$ is not diagonalizable. Hence, we can see that A and B are both diagonalizable, while AB is not diagonalizable.

(c) The statement is true. If $A^2 = A$, then $A(A - I) = 0$. It follows that the minimal polynomial of A is either x (if $A = 0$), or $(x - 1)$ (if $A = I$), or $x(x - 1)$. In any case, μ_A has no repeated roots, and thus A is diagonalizable.

(d) The statement is not true. A counterexample is

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ with } A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

The matrix A is a Jordan block of size 2 associated with eigenvalue 0, so A is not diagonalizable. The matrix A^2 is diagonal, so A^2 is diagonalizable. Hence, we can see that A is not diagonalizable, while A^2 is diagonalizable.

Problem 4

Let A, B, and C represent three real $n \times n$ matrices, where A and B be symmetric positive definite (spd) and C be invertible. Prove that each of the following is spd.   [5 points each]

(a) A^{-1}

(b) $A + B$

(c) $C^T A C$

(d) $A^{-1} - (A + B)^{-1}$

Solution: We use the definition and property that a real symmetric matrix A is positive definite if and only if $x^T A x > 0$ for all real n-dimensional vectors $x \neq 0$, or equivalently, if all its eigenvalues are real and positive.

(a) Since A is symmetric, we have
$$(A^{-1})^T = (A^{-1})^T A A^{-1} = (A^T A^{-1})^T A^{-1} = (A A^{-1})^T A^{-1} = A^{-1},$$
and thus A^{-1} is symmetric. Moreover, if A is positive definite, then all of its eigenvalues are positive, and if $Ax = \lambda x$, then $A^{-1}x = \lambda^{-1}x$, so A^{-1} is positive definite as well.

(b) Since A and B are symmetric, clearly, $A + B$ is symmetric. Moreover, we have
$$x^T(A + B)x = x^T A x + x^T B x,$$
and since A and B are spd, it follows that $x^T A x > 0$ and $x^T B x > 0$ for any $x \neq 0$. This implies that also $x^T(A + B)x > 0$, and thus $A + B$ is positive definite as well.

(c) Since A is symmetric, we have
$$(C^T A C)^T = C^T A^T C = C^T A C,$$
and thus $C^T A C$ is symmetric. Then let $x \neq 0$, so that $Cx \neq 0$ because C is invertible. It follows that
$$x^T C^T A C x = (Cx)^T A(Cx) > 0,$$
and thus $C^T A C$ is positive definite as well.

(d) First, it is clear that $A^{-1} - (A + B)^{-1}$ is symmetric if A and B are symmetric, because inverses and sums of symmetric matrices are symmetric; also compare parts (a) and (b). Second, from
$$A^{-1} - (A + B)^{-1} = \left(A^{-1}(A + B) - I\right)(A + B)^{-1} = A^{-1}B(A + B)^{-1}$$
and
$$\left(A^{-1}B(A + B)^{-1}\right)^{-1} = (A + B)B^{-1}A = AB^{-1}A + A,$$
it follows that the inverse of $A^{-1} - (A + B)^{-1}$ is spd from part (a) (applied to B), part (b) (applied to $AB^{-1}A$ and A), and part (c) (applied to $AB^{-1}A$). Hence, $A^{-1} - (A + B)^{-1}$ is spd again from part (a).

Problem 5

Let $\mathcal{P}_2[0,2]$ represent the set of polynomials with real coefficients and of degree less than or equal to 2, defined on $[0,2]$. For $p = (p(t)) \in \mathcal{P}_2$ and $q = (q(t)) \in \mathcal{P}_2$, define

$$\langle p, q \rangle := p(0)q(0) + p(1)q(1) + p(2)q(2).$$

(a) Verify that $\langle p, q \rangle$ is an inner product. [4 points]

(b) Let T represent the linear transformation that maps an element $p \in \mathcal{P}_2$ to the closest element of the span of the polynomials 1 and t in the sense of the norm associated with the inner product. Find the matrix A of T in the standard basis of $\mathcal{P}_2$.
(Note: the standard basis of $\mathcal{P}_2$ is $\{1, t, t^2\}$.) [10 points]

(c) Is A symmetric? Is T self-adjoint? Do these facts contradict each other? [3 points]

(d) Find the minimal polynomial of T. [3 points]

Solution:

(a) We need to show that $\langle p, q \rangle$ is symmetric, bilinear, and positive definite. Let p, q, and r be three elements in $\mathcal{P}_2$, and let α and β be two real numbers. Bilinearity (*i.e.*, linearity with respect to first argument: $\langle \lambda p, q \rangle = \lambda \langle p, q \rangle$ and $\langle p + q, r \rangle = \langle p, r \rangle + \langle q, r \rangle$; and linearity with respect to second argument: $\langle p, \lambda q \rangle = \lambda \langle p, q \rangle$ and $\langle p, q + r \rangle = \langle p, q \rangle + \langle p, r \rangle$), symmetry (*i.e.*, $\langle p, q \rangle = \langle q, p \rangle$), and positiveness (*i.e.*, $\langle p, p \rangle \geq 0$) are mechanical.

Definiteness (*i.e.*, $\langle p, p \rangle = 0 \Rightarrow p = 0$) requires some attention, so let $p \in \mathcal{P}_2$ such that $\langle p, p \rangle = 0$. It follows that $p(0)^2 + p(1)^2 + p(2)^2 = 0$, so $p(0) = p(1) = p(2) = 0$. But we know that the only polynomial of degree less than or equal to 2 that has three roots is the zero polynomial, so $p = 0$.

(b) We understand that T is the orthogonal projection onto the subspace spanned by 1 and t. To find the matrix of T in the standard basis, let us apply T to 1, t, and t^2. It is clear that $T(1) = 1$, and that $T(t) = t$. Now we need to compute $T(t^2)$. So we need to compute the orthogonal projection of t^2 onto the subspace spanned by 1 and t. Let us find an orthogonal basis for the subspace spanned by 1 and t. Using the Gram-Schmidt algorithm, we get $v_1(t) = 1$ and

$$v_2(t) = t - \frac{\langle t, 1 \rangle}{\langle 1, 1 \rangle} 1 = t - 1.$$

Hence, $\{v_1, v_2\}$ is an orthogonal basis for the subspace spanned by 1 and t. Using this orthogonal basis, we can now perform the orthogonal projection of t^2 onto 1 and t:

$$T(t^2) = \frac{\langle t^2, 1 \rangle}{\langle 1, 1 \rangle} 1 + \frac{\langle t^2, t - 1 \rangle}{\langle t - 1, t - 1 \rangle} (t - 1) = \frac{5}{3} + 2(t - 1) = -\frac{1}{3} + 2t.$$

Thus, the standard basis $\{1, t, t^2\}$ is mapped to $\{1, t, -1/3 + 2t\}$. In coordinate vectors, $(1, 0, 0)$ is mapped into $(1, 0, 0)$, $(0, 1, 0)$ is mapped into $(0, 1, 0)$, and $(0, 0, 1)$ is mapped to $(-1/3, 2, 0)$, so the transformation matrix is

$$A = \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

(c) The matrix A is not symmetric. The transformation is self-adjoint being an orthogonal projection: since $\langle Tp, q - Tq \rangle = 0$ and $\langle Tp - p, Tq \rangle = 0$, we have that $\langle Tp, q \rangle = \langle Tp, Tq \rangle = \langle p, Tq \rangle$. The matrix A of the transformation T is given in the basis $\{1, t, t^2\}$, which is not an orthogonal basis, so the facts that $x^T A y \neq x^T A^T y$ (matrix is not symmetric) and that $\langle Tp, q \rangle = \langle p, Tq \rangle$ (operator is self-adjoint) do not contradict each other.

(d) Since T is a projection, we know that $T^2 - T = 0$. Moreover, $T \neq I$ (so the minimal polynomial is not $x - 1$), $T \neq 0$ (so the minimal polynomial is not x), and thus the minimal polynomial is $\mu_T(x) = x^2 - x$.

Alternative Computation of the Transformation Matrix (using optimization): Given a quadratic polynomial $p(t) = a_0 + a_1 t + a_2 t^2$, the orthogonal mapping onto (the "closest" element in) the span of the polynomials 1 and t can be interpreted as the problem to find the matrix A that maps the vector (a_0, a_1, a_2) to a vector (b_0, b_1) that corresponds to the "best" linear fit $f(t) = b_0 + b_1 t$, in the sense that b_0 and b_1 minimize the norm

$$\|p - q\|^2 = (p - q, p - q) = (a_0 - b_0)^2 + (a_0 + a_1 + a_2 - b_0 - b_1)^2 + (a_0 + 2a_1 + 4a_2 - b_0 - 2b_1)^2.$$

This can be formulated as unconstrained (convex quadratic) norm minimization problem and solved by computing and solving for b_0 and b_1 in the vanishing gradient of $\|p - q\|^2$:

$$\partial/\partial b_0 = 0 \quad \Rightarrow \quad (a_0 - b_0) + (a_0 + a_1 + a_2 - b_0 - b_1) + (a_0 + 2a_1 + 4a_2 - b_0 - 2b_1)$$
$$= \quad 3a_0 + 3a_1 + 5a_2 - 3b_0 - 3b_1 = 0;$$
$$\partial/\partial b_1 = 0 \quad \Rightarrow \quad (a_0 + a_1 + a_2 - b_0 - b_1) + 2(a_0 + 2a_1 + 4a_2 - b_0 - 2b_1)$$
$$= \quad 3a_0 + 5a_1 + 9a_2 - 3b_0 - 5b_1 = 0.$$

Multiplying these equations by $+5$ and -3, or by -1 and $+1$, and adding, we obtain

$$6a_0 - 2a_2 - 6b_0 = 0 \quad \Rightarrow \quad b_0 = a_0 - (1/3)a_2;$$
$$2a_1 + 4a_2 - 2b_1 = 0 \quad \Rightarrow \quad b_1 = a_1 + 2a_2;$$

and thus $A = \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & 2 \end{pmatrix}$. (We may also add a zero row for the third component.)

Problem 6

Suppose that A is an $m \times n$ matrix and B is an $n \times m$ matrix, and write I_m for the $m \times m$ identity matrix. Show that if $I_m - AB$ is invertible, then so is $I_n - BA$.   [20 points]

Solution: To show that $I_n - BA$ is invertible, it suffices to show that $I_n - BA$ has the trivial nullspace, so let $x \in \mathbb{R}^n$ such that $x - BAx = 0$. It follows that $BAx = x$, so $AB(Ax) = Ax$, and thus Ax is in the nullspace of $I_m - AB$. In particular, because $I_m - AB$ is invertible, that nullspace is trivial, so that $Ax = 0$ and thus $x = BAx = 0$.

Alternative Solution (using eigenvalues):   It is known that the eigenvalues of BA and AB are the same at the exception of the possible 0 eigenvalue. Then $I_n - BA$ is invertible if and only if 0 is not an eigenvalue of $I_m - BA$, which means that 1 is not an eigenvalue of BA, which means that 1 is not an eigenvalue of AB, which means that 0 is not an eigenvalue of $I_n - AB$, and thus, if and only if $I_m - AB$ is invertible.

To show that two product matrices AB and BA have the same nonzero eigenvalues, let λ be an eigenvalue of AB with eigenvector $x \neq 0$, so that $ABx = \lambda x$ and $BABx = \lambda Bx$. It follows that λ is also an eigenvalue of BA, if Bx is nonzero; otherwise, if $Bx = 0$, then $\lambda x = 0$ and thus $\lambda = 0$ because $x \neq 0$. Hence, λ is an eigenvalue of AB if and only if it is also an eigenvalue of BA or zero.

Alternative Solution (by construction of the inverse):   Let $I_m - AB$ be invertible and consider the linear system $(I_n - BA)x = b$ for any $b \in \mathbb{R}^n$. To show that $I_n - BA$ is invertible, we may show that this system has a unique solution $x \in \mathbb{R}^m$. Now note that

$$(I_n - BA)x = x - BAx = b$$
$$\Rightarrow \quad Ax - ABAx = Ab$$
$$\Rightarrow \quad (I_m - AB)Ax = Ab$$
$$\Rightarrow \quad Ax = (I_m - AB)^{-1}Ab$$

and thus, from $x - BAx = b$, that

$$x = b + BAx = b + B(I_m - AB)^{-1}Ab = (I_n + B(I_m - AB)^{-1}A)b.$$

This suggests that $(I_n - BA)^{-1} = I_n + B(I_m - AB)^{-1}A$, which can be confirmed as follows:

$$(I_n - BA)\left(I_n + B(I_m - AB)^{-1}A\right) = I_n - BA + B(I_m - AB)(I_m - AB)^{-1}A = I_n;$$
$$\left(I_n + B(I_m - AB)^{-1}A\right)(I_n - BA) = I_n - BA + B(I_m - AB)^{-1}(I_m - AB)A = I_n.$$

Alternative Derivation of the Inverse   Alternatively known as the matrix inversion lemma or the (Sherman-Morrison-)Woodbury matrix identity, if $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{m \times m}$, and $D \in \mathbb{R}^{n \times n}$, then

$$(D + BCA)^{-1} = D^{-1} - D^{-1}B\left(C^{-1} + AD^{-1}B\right)^{-1}AD^{-1}.$$

Hence, the same inverse as before follows immediately with $C = -I_m$ and $D = I_n$, because

$$(I_n - BA)^{-1} = I_n - B\left(-I_m + AB\right)^{-1}A = I_n + B\left(I_m - AB\right)^{-1}A.$$