MATH 423

Linear Algebra II

Lecture 15: Inverse matrix (continued). Transpose of a matrix.

Elementary row operations for matrices:

(1) to interchange two rows;

(2) to multiply a row by a nonzero scalar;

(3) to add the ith row multiplied by some scalar r to the jth row.

Similarly, we define three types of **elementary column operations**.

- Any elementary row operation σ on matrices with n rows can be simulated as left multiplication by a certain $n \times n$ matrix E_σ (called **elementary**).
- The elementary matrix E_σ is obtained by applying the operation σ to the identity matrix.
- Any elementary column operation can be simulated as right multiplication by a certain elementary matrix.
- Elementary matrices are invertible.

General results on inverse matrices

Theorem 1 Given a square matrix A, the following are equivalent:

(i) A is invertible;
(ii) $\mathbf{x} = \mathbf{0}$ is the only solution of the matrix equation $A\mathbf{x} = \mathbf{0}$ (where $\mathbf{x}$ and $\mathbf{0}$ are column vectors).

Theorem 2 For any $n \times n$ matrices A and B,

$$BA = I \iff AB = I \iff B = A^{-1}.$$

Theorem 3 Suppose that a sequence of elementary row operations converts a matrix A into the identity matrix. Then A is invertible. Moreover, the same sequence of operations converts the identity matrix into the inverse matrix A^{-1}.

Theorem 1 Given an $n \times n$ matrix A, the following are equivalent: **(i)** A is invertible; **(ii)** $\mathbf{x} = \mathbf{0}$ is the only solution of the matrix equation $A\mathbf{x} = \mathbf{0}$.

Proof: **(i)** $\Longrightarrow$ **(ii)** Assume A is invertible. Take any column vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$. Then $A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{0}$. We have $A^{-1}(A\mathbf{x}) = (A^{-1}A)\mathbf{x} = I\mathbf{x} = \mathbf{x}$ and $A^{-1}\mathbf{0} = \mathbf{0}$. Hence $\mathbf{x} = \mathbf{0}$.

(ii) $\Longrightarrow$ **(i)** Assume $\mathbf{x} = \mathbf{0}$ is the only solution of the matrix equation $A\mathbf{x} = \mathbf{0}$. Consider a linear operator $L_A : \mathbb{F}^n \to \mathbb{F}^n$ given by $L_A(\mathbf{x}) = A\mathbf{x}$. By assumption, the null-space $\mathcal{N}(L_A)$ is trivial. It follows that L_A is one-to-one. By the Dimension Theorem, $\dim \mathcal{R}(L_A) + \dim \mathcal{N}(L_A) = \dim \mathbb{F}^n = n$. Then $\dim \mathcal{R}(L_A) = n$, which implies that $\mathcal{R}(L_A) = \mathbb{F}^n$. That is, L_A is onto. Thus L_A is an invertible mapping.

The inverse L_A^{-1} is also linear. Hence $L_A^{-1}(\mathbf{x}) = B\mathbf{x}$ for some $n \times n$ matrix B and any column vector $\mathbf{x} \in \mathbb{F}^n$. Clearly, $L_A^{-1}(L_A(\mathbf{x})) = \mathbf{x} = L_A(L_A^{-1}(\mathbf{x}))$, i.e., $BA\mathbf{x} = \mathbf{x} = AB\mathbf{x}$, for all $\mathbf{x}$. It follows that $BA = I = AB$. Thus $B = A^{-1}$.

Theorem 2 For any $n \times n$ matrices A and B,

$$BA = I \iff AB = I \iff B = A^{-1}.$$

Proof: $[BA = I \implies B = A^{-1}]$ Assume $BA = I$. Take any column vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$. Then $B(A\mathbf{x}) = B\mathbf{0}$. We have $B(A\mathbf{x}) = (BA)\mathbf{x} = I\mathbf{x} = \mathbf{x}$ and $B\mathbf{0} = \mathbf{0}$. Hence $\mathbf{x} = \mathbf{0}$. By Theorem 1, A is invertible. Then $BA = I \implies (BA)A^{-1} = IA^{-1}$ $\implies B = A^{-1}$.

$[AB = I \implies B = A^{-1}]$ Assume $AB = I$. By the above B is invertible and $A = B^{-1}$. The latter is equivalent to $B = A^{-1}$.

Proof of Theorem 3

Assume that a square matrix A can be converted to the identity matrix by a sequence of elementary row operations. Then

$$E_k E_{k-1} \dots E_2 E_1 A = I,$$

where $E_1, E_2, \dots, E_k$ are elementary matrices simulating those operations.

Applying the same sequence of operations to the identity matrix, we obtain the matrix

$$B = E_k E_{k-1} \dots E_2 E_1 I = E_k E_{k-1} \dots E_2 E_1.$$

Thus $BA = I$, which, by Theorem 2, implies that $B = A^{-1}$.

Let V denote the set of all solutions of a differential equation $u'''(x) - 2u''(x) + u'(x) = 0, \ x \in \mathbb{R}$.

The set V is a subspace of $C^3(\mathbb{R})$ since it is the null-space of a linear diffrential oprator $L : C^3(\mathbb{R}) \to C(\mathbb{R})$ given by $Lu = u''' - 2u'' + u'$.

According to the theory of differential equations, the initial value problem

$$\begin{cases} u'''(x) - 2u''(x) + u'(x) = 0, \quad x \in \mathbb{R}, \\ u(0) = b_0, \\ u'(0) = b_1, \\ u''(0) = b_2 \end{cases}$$

has a unique solution for any $b_0, b_1, b_2 \in \mathbb{R}$. In other words, a linear mapping $J : V \to \mathbb{R}^3$, given by $J(u) = \big(u(0), u'(0), u''(0)\big)$, is one-to-one and onto, i.e., invertible.

Problem. Find the inverse transformation J^{-1}.

We know from the previous lecture that functions $u_1(x) = 1$, $u_2(x) = e^x$, and $u_3(x) = xe^x$ form a basis for V. Let α denote this basis and β denote the standard basis for $\mathbb{R}^3$. We are going to find the matrix $[J]_\alpha^\beta$.

$u_1'(x) = u_1''(x) = 0$, $u_2'(x) = u_2''(x) = e^x$,
$u_3'(x) = xe^x + e^x$, $u_3''(x) = xe^x + 2e^x$.

$$[J]_\alpha^\beta = \begin{pmatrix} u_1(0) & u_2(0) & u_3(0) \\ u_1'(0) & u_2'(0) & u_3'(0) \\ u_1''(0) & u_2''(0) & u_3''(0) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Let $A = [J]_\alpha^\beta$. Then the matrix $[J^{-1}]_\beta^\alpha$ is A^{-1}.

A convenient way to compute the inverse matrix A^{-1} is to merge the matrices A and I into one 3×6 matrix $(A \,|\, I)$ and apply elementary row operations to this new matrix. The goal is to get a matrix of the form $(I \,|\, B)$, then $B = A^{-1}$.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \qquad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$(A \,|\, I) = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right)$$

$$(A\,|\,I) = \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 0 & 0 & 1 \end{array}\right) \to$$

$$\left(\begin{array}{ccc|rrr} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|rrr} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|rrr} 1 & 0 & 0 & 1 & -2 & 1 \\ 0 & 1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 1 & 0 & -1 & 1 \end{array}\right) = (I\,|\,A^{-1}).$$

It follows that $J^{-1}(\mathbf{e}_1) = u_1$, $J^{-1}(\mathbf{e}_2) = -2u_1 + 2u_2 - u_3$, $J^{-1}(\mathbf{e}_3) = u_1 - u_2 + u_3$.

For any vector $(y_1, y_2, y_3) \in \mathbb{R}^3$ we have $J^{-1}(y_1, y_2, y_3) = f$, where $f = y_1 u_1 + y_2(-2u_1 + 2u_2 - u_3) + y_3(u_1 - u_2 + u_3)$ so that $f(x) = (y_1 - 2y_2 + y_3) + (2y_2 - y_3)e^x + (-y_2 + y_3)xe^x$.

Transpose of a matrix

Definition. Given a matrix A, the **transpose** of A, denoted A^t, is the matrix whose rows are columns of A (and whose columns are rows of A). That is, if $A = (a_{ij})$ then $A^t = (b_{ij})$, where $b_{ij} = a_{ji}$.

Examples. $$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix},$$

$$\begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}^t = (7, 8, 9), \qquad \begin{pmatrix} 4 & 7 \\ 7 & 0 \end{pmatrix}^t = \begin{pmatrix} 4 & 7 \\ 7 & 0 \end{pmatrix}.$$

Properties of transposes:

- $(A^t)^t = A$
- $(A + B)^t = A^t + B^t$
- $(rA)^t = rA^t$
- $(AB)^t = B^t A^t$
- $(A_1 A_2 \ldots A_k)^t = A_k^t \ldots A_2^t A_1^t$
- $(A^{-1})^t = (A^t)^{-1}$