Lecture 9: Multiplication and linear transformations

February 8, 2024

Gaussian elimination for inverse calculation

Example $A = \begin{bmatrix} 1 & -1 & 1 & 4 \\ 1 & 0 & 2 & 2 \\ 3 & -3 & 4 & 8 \\ 0 & -2 & -2 & 5 \end{bmatrix}$. Find A^{-1}.

$$\left[\begin{array}{cccc|cccc} 1 & -1 & 1 & 4 & 1 & 0 & 0 & 0 \\ 1 & 0 & 2 & 2 & 0 & 1 & 0 & 0 \\ 3 & -3 & 4 & 8 & 0 & 0 & 1 & 0 \\ 0 & -2 & -2 & 5 & 0 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 26 & -19 & -2 & -10 \\ 0 & 1 & 0 & 0 & 6 & -3 & -1 & -2 \\ 0 & 0 & 1 & 0 & -11 & 8 & 1 & 4 \\ 0 & 0 & 0 & 1 & -2 & 2 & 0 & 1 \end{array}\right]$$

Conclusion $A^{-1} = \begin{bmatrix} 26 & -19 & -2 & -10 \\ 6 & -3 & -1 & -2 \\ -11 & 8 & 1 & 4 \\ -2 & 2 & 0 & 1 \end{bmatrix}$. Check this!

If I_n is written in the first n columns of the RREF of $[A|I_n]$, the last n columns comprise A^{-1}.

Why did that work?

1. We were solving four linear systems simultaneously. All four had the same coefficient matrix A, their right hand sides were respectively ae_1, e_2, e_3, e_4.
2. The four leading 1's in the RREF mean that in each system (and any other with coefficient matrix A), there is a unique solution. Those unique solutions are respectively written in the last four columns. So Column 5 of the RREF is the unique column vector v with $Av = e_1$. This is the first column of A^{-1}.
3. If A didn't have an inverse, what would have happened?
If A has no inverse, at least one of the four systems is inconsistent. In the row reduction, we encounter a row with 0 in the first four positions, but not in all the last four.

Elementary row operations as multiplication by matrices

Let A be (for example) a 4×4 matrix. Applying an elementary row operation (ERO) to A means multiplying A on the left by another matrix.

ERO Type 1 Adding $4\times$ Row 2 to Row 3 means multiplying A on the left by

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

ERO Type 2 Swapping Rows 2 and 4 means multiplying A on the left by

$$E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

ERO Type 3 Multiplying Row 3 by 5 means multiplying A on the left by

$$E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

To see this, think about the effect of on the rows of A of left multiplication by E_1, E_2, E_3. Each of these E_i is an elementary matrix.

Gaussian elimination and factorization

When we apply Gaussian elimination to reduce a matrix A to RREF, we are identifying a sequence $E_1, E_2, \dots, E_k$ with

$$RREF(A) = E_k \dots E_2 E_1 A.$$

An $n \times n$ elementary matrix can differ from I_n in one of the following ways

- By having one non-zero entry away from the main diagonal.
- By having one non-zero entry on the main diagonal that is not 1.
- By swapping the columns of two of the 1s in I_n.

Exercise Show that the inverse of an elementary matrix is an elementary matrix.

With this in mind, we can rewrite the above equation as

$$A = E_1^{-1} E_2^{-1} \dots E_k^{-1} \, RREF(A).$$

Any matrix A can be factorized as $A = FB$, where F is a product of elementary matrices and B is a RREF. This factorization is very useful in numerical/computational mathematics.

Linear Transformations

Linear transformations are the primary functions between vector spaces that are of interest in linear algebra. They are special because they cooperate with the algebraic structure.

Definition Let m and n be positive integers. A linear tranformation T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a function $T : \mathbb{R}^n \to \mathbb{R}^m$ that satisfies

- $T(u + v) = T(u) + T(v)$, and
- $T(\lambda v) = \lambda T(v)$,

for all u and v in $\mathbb{R}^n$, and all scalars $\lambda \in \mathbb{R}$.

The Matrix of a Linear Transformation

Suppose that $T : \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation, with

$$T\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}2\\-3\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix} = \begin{bmatrix}1\\4\end{bmatrix}, T\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}-6\\7\end{bmatrix}$$

Then for the vector in $\mathbb{R}^3$ with any entries a, b, c

$$T\begin{bmatrix}a\\b\\c\end{bmatrix} = aT\begin{bmatrix}1\\0\\0\end{bmatrix} + bT\begin{bmatrix}0\\1\\0\end{bmatrix} + cT\begin{bmatrix}0\\0\\1\end{bmatrix} = \underbrace{\begin{bmatrix}2 & 1 & -6\\-3 & 4 & 7\end{bmatrix}}_{M_T}\begin{bmatrix}a\\b\\c\end{bmatrix}.$$

and the 2×3 matrix M_T is called the (standard) matrix of A.

The matrix of a linear transformation

- A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is represented by a $m \times n$ matrix M_T. The columns of M_T are the images under T of the standard basis vectors $e_1, \dots, e_n$.
- If v is any vector in $\mathbb{R}^n$, we can calculate $T(v)$ by multilpying the column vector v on the left by the matrix M_T. Matrix-vector multiplication is evaluating linear transformations.
- On the other hand, if A is any $m \times n$ matrix, then A determines a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ by $v \to Av$, for $v \in \mathbb{R}^n$. So, in a sense, matrices are linear transformations.
- Examples of linear transformations include rotations, reflections and scaling, but not translations.
- If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then in order to evaluate T at any point/vector, we only need mn pieces of information, just the m coordinates of each of the n images of the standard basis vectors. This is very different for example from continuous functions from $\mathbb{R}$ to $\mathbb{R}$ - we cannot know all about them just by knowing their values at a few points.

Matrix multplication is composition

Suppose that $T : \mathbb{R}^n \to \mathbb{R}^p$ and $S : \mathbb{R}^p \to \mathbb{R}^m$ are linear transformations. Then $S \circ T$ (S after T) is the linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined for $v \in \mathbb{R}^n$ by

$$S \circ T(v) = S(T(v)).$$

Question How does the $(m \times n)$ matrix $M_{S \circ T}$ of $S \circ T$ depend on the $(m \times p)$ matrix M_S of S and the $(p \times n)$ matrix M_T of T?
To answer this we have to think about the definition of $M_{S \circ T}$.

- Its first column has the coordinates of $S \circ T(e_1) = S(T(e_1))$.
- $T(e_1)$ is the first column of M_T.
- Then $S(T(e_1))$ is the matrix-vector product M_S[first column of M_T]. This is the first column of the matrix product $M_S M_T$.
- Same for all the other columns: the conclusion is $M_{S \circ T} = M_S M_T$.

Matrix multiplication is composition of linear transformations.

Corollary Matrix multiplication is associative.