Elementary Matrices and Matrix Invertibility

Christian B. Hughes

BIE-LA1 - Winter 2025

Contents

1 Matrix Invertibility **1**
1.1 Matrices as Linear Maps 1
1.2 Matrix Invertibility 2
1.3 Exercises 3

2 Elementary Matrices **3**
2.1 Types of Elementary Matrices 3
2.2 Gaussian Elimination as Matrix Multiplication 3
2.3 Computation of Matrix Inverses 6
2.4 Exercises 7

1 Matrix Invertibility

1.1 Matrices as Linear Maps

We have employed matrices as tools in our investigation of properties of finite-dimensional vector spaces for a few weeks now, but almost exclusively in that capacity; we have yet to explore the structural properties of matrices themselves in depth. We do so now, beginning with the necessary elements of the general theory of linear maps on finite-dimensional vector spaces, which will inform our knowledge of matrix properties.

Let V and U be finite dimensional vector spaces over the field $\mathbb{F}$, and let $T : V \to U$ be a map satisfying the following properties:

1. For all $v, u \in V$, $T(v + u) = T(u) + T(v)$.
2. For all $c \in \mathbb{F}$, $T(cv) = cT(v)$.

Recall that, for any $n \times m$ matrix M over the field $\mathbb{F}$, and for any vector $v \in \mathbb{F}^m$, we have $Mv \in \mathbb{F}^n$. One can thus identify M as a function from $\mathbb{F}^m$ to $\mathbb{F}^n$ Furthermore, any $m \times n$ matrix of this form satisfies the two properties above (this is left as an exercise), and thus matrices are not only functions, but a special kind of function: a linear map.

Remark 1.1. *One might wonder why this "linearity" idea seems to pervade the subject of "linear algebra." Aside from being the namesake of the subject, linear maps are more or less core to all human technology. Continuous objects are often extremely difficult or impossible to model practically, and thus we resort to local linear approximations (linear functions often described by matrices) of phenomena we wish to model. Want to understand the local effects of the EM field on a complicated machine? Derivatives and a matrix. Want to build a bridge? Derivatives and a matrix. More on this in LA2, MA1, and MA2.*

1.2 Matrix Invertibility

Given that matrices are special kinds of functions on vector spaces, it is natural to ask when such functions are invertible. Recall that a general function $f : A \to B$ is invertible if and only if there exists a function $g : B \to A$ such that $g \circ f = id_A$ for all $x \in A$, where id_A denotes the identity function on A.

We will restrict ourselves to $M_n(\mathbb{F})$ for the time being. Viewing matrix multiplication as function composition (a notion to which we will return to later), we would like to explicitly provide a similar notion of invertibility in this context.

A matrix $A \in M_n(\mathbb{F})$ is *invertible* if and only if there exists a (not necessarily distinct) matrix $B \in M_n(\mathbb{F})$ such that $AB = BA = I_n$, where I_n denotes the $n \times n$ matrix with the multiplicative identity of $\mathbb{F}$ along its main diagonal, with the additive identity of $\mathbb{F}$ in every other entry. A matrix A with an inverse is called invertible, and its inverse is denoted by A^{-1}.

Given a matrix $A \in M_n(\mathbb{F})$, if there exists a matrix B such that $BA = I_n$, we call B a *left inverse* for A. Similarly, if there exists a matrix B such that $AB = I_n$, we say that B is a *right inverse* for A.

We can deduce the following two propositions from these definitions, which are provided without proof.

Proposition 1.1. *If a matrix is invertible, then it is a square matrix.*

Proposition 1.2. *If a matrix is invertible on the right (resp. left), then the matrix is invertible on the left (resp. right) by the same inverse.*

Recall that a function is invertible iff there is an exact one-to-one correspondence between its domain and codomain. However, by the linearity of matrices, we are asking something more of our maps between vector spaces: that they somehow "preserve" the structure of the vector spaces.

Thus, the intuition for Proposition 1.1 is that, despite the fact that $\mathrm{Card}(\mathbb{R}) = \mathrm{Card}(\mathbb{R}^2)$[1], and hence that there is a bijection between $\mathbb{R}$ and $\mathbb{R}^2$, asking for a *structure-preserving* bijection between these sets as vector spaces puts too many constraints on such a bijection, and is not possible; we need the same number of dimensions to work with in both the domain and codomain to obtain a linear map (matrix) between them which is invertible.

[1] Here, "Card" denotes *cardinality*.

As a natural consequence of the definition of invertibility, we have the following proposition.

Proposition 1.3. *Let A and B be invertible matrices. Then $(AB)^{-1} = B^{-1}A^{-1}$.*

Proof. For AB, we have $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$. □

1.3 Exercises

1. Find the formula for the inverse of a product of $n \in \mathbb{N}$ invertible matrices.
2. Provide an example of a non-invertible matrix.
3. Would the property of non-invertibility of $n \times m$ matrices, $n \times m$, change with a particular choice of field? If so, which field? If not, why not?
4. Provide an example of an invertible matrix.
5. ($\star$) What if we dropped the word "field" in Question 3 and replaced it with some other kind of structure?

2 Elementary Matrices

Until now, we have viewed matrices in Gaussian elimination as a convenient tool, but nothing more. However, in an effort to better understand why Gaussian elimination works, we would like to link the process of Gaussian elimination itself to the notions we have developed so far. We do this via the multiplication of *elementary matrices*.

2.1 Types of Elementary Matrices

Recall that we have three fundamental row operations in Gaussian elimination:

1. Multiplication of a row by a (nonzero) scalar;
2. Interchange of rows;
3. Addition of one row to another.

We will now associate matrices to each of these operations.

2.2 Gaussian Elimination as Matrix Multiplication

Suppose we have the $n \times n$ matrix M over the field $\mathbb{F}$, and we would like to multiply the i^{th} row, $i < n$, of M by $c \in \mathbb{F}$. Concretely, we can write

$$M = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{i,1} & \cdots & a_{i,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}.$$

Pay special attention to the indices of the entries of M; despite the rectangular appearance in its formatting here, M is square.

Now consider the matrix E_1:

$$E_1 = \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & c & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

where c occurs in the i^{th} row and i^{th} column of E_1, with ones along the rest of the main-diagonal entries, with zeros everywhere else. We obtain

$$E_1 M = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ ca_{i,1} & \cdots & ca_{i,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}.$$

Consequently, matrices in the form of E_1, when multiplied by a matrix from the left, correspond to row-scalar multiplication operation of Gaussian elimination.

Now suppose we would like to exchange two rows m_1 and m_2 of M instead. We consider the $n \times n$ matrix E_2 such that the following equation holds for the entries $e_{i,j}$ of E, $1 \le i, j \le n$:

$$e_{i,j} = \begin{cases} 1 \text{ if } i = j \text{ and } i \notin \{m_1, m_2\}, \\ 1 \text{ if } (i,j) = (m_1, m_2) \text{ or } (i,j) = (m_2, m_1), \\ 0 \text{ otherwise.} \end{cases}$$

To make this intuitive, we provide an example.

Example 2.1. *Suppose we have the matrix*

$$M = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix}.$$

We would like to swap rows 1 *and* 3*. Employing the definition of* E_2*, we find*

$$E_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Thus,

$$E_2 M = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} = \begin{pmatrix} a_{3,1} & a_{3,2} & a_{3,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{1,1} & a_{1,2} & a_{1,3} \end{pmatrix}.$$

Finally, we consider the elementary matrix E_3 corresponding to the Gaussian elimination step of adding a (nonzero) multiple of one row to another. Suppose we would like to represent the row operation $R_{m_1} \leftarrow R_{m_1} + cR_{m_2}$, where c is some scalar in the underlying field. Define E_3 as the $n \times n$ matrix with entries $e_{i,j}$, $1 \le i, j \le n$, such that

$$e_{i,j} = \begin{cases} 1 \text{ if } i = j, \\ c \text{ if } (i,j) = (m_1, m_2), \\ 0 \text{ otherwise.} \end{cases}$$

Informally, this matrix is the identity everywhere, except for (m_1, m_2), in which we place c.

Example 2.2. *Consider the matrix*

$$M = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}.$$

We would like to perform the Gaussian elimination operation $R_2 \leftarrow R_2 + 3R_1$. We form the matrix

$$E_3 = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}.$$

Multiplying these matrices, we find that

$$E_3 M = \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} \\ 3a_{1,1} + a_{2,1} & 3a_{1,2} + a_{2,2} \end{pmatrix},$$

as desired.

The following proposition will prove useful for explicitly computing matrix inverses.

Proposition 2.1. *The elementary matrices E_1, E_2, E_3* [2] *are invertible.*

Proof. See Exercise 2.4.3. □

[2] Note that the notation E_1, E_2, and E_3 for the three elementary matrices is not universally standard.

2.3 Computation of Matrix Inverses

We now leverage our results concerning elementary matrices to obtain a method for computing matrix inverses. First, suppose that we have some $n \times n$ matrix A, and that A can be reduced to the identity matrix I in p Gaussian elimination steps. Denote the i^{th} elementary matrix in this sequence of steps by E_i. Then

$$I = E_p E_{p-1} \cdots E_1 A.$$

Taking the inverse of both sides, we have

$$(I)^{-1} = (E_p E_{p-1} \cdots E_1 A)^{-1} = A^{-1} E_1^{-1} \cdots E_{p-1}^{-1} E_p^{-1}.$$

Solving for A^{-1} by multiplying both sides by the p elementary matrices, we obtain

$$E_p E_{p-1} \cdots E_1 = E_p E_{p-1} \cdots E_1 I = A^{-1}.$$

Consequently, we observe that, if A can be row-reduced to the identity matrix, A^{-1} can be recovered by multiplying the elementary matrices corresponding to the Gaussian elimination steps that row-reduced A.

This leads us to the following proposition

Proposition 2.2. *A matrix is invertible if and only if it can be row-reduced to the identity matrix.*

Proof. See Exercise 2.4.4. $\square$

In practice, when row reduce an augmented matrix, we apply the row operations to the augmented column vector as well. Thus, if we augment our matrix with the entire identity matrix, all row operations used to reduce our augmented matrix to the identity will be applied to the identity matrix, recovering our matrix inverse.

Example 2.3. *Consider the matrix*

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

We augment with the identity matrix and apply Gaussian elimination:

$$\left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 3 & 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

$$R_1 \leftrightarrow R_2 : \quad \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$R_2 \leftarrow R_2 - 2R_1: \quad \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & -5 & -3 & 1 & -2 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$R_3 \leftarrow R_3 - R_1: \quad \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & -5 & -3 & 1 & -2 & 0 \\ 0 & -3 & -2 & 0 & -1 & 1 \end{array}\right)$$

$$R_3 \leftarrow R_3 - \frac{3}{5}R_2: \quad \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & -5 & -3 & 1 & -2 & 0 \\ 0 & 0 & -\frac{1}{5} & -\frac{3}{5} & \frac{1}{5} & 1 \end{array}\right)$$

$$R_3 \leftarrow -5R_3: \quad \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 0 & 1 & 0 \\ 0 & -5 & -3 & 1 & -2 & 0 \\ 0 & 0 & 1 & 3 & -1 & -5 \end{array}\right)$$

$$R_1 \leftarrow R_1 - 2R_3: \quad \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & -6 & 3 & 10 \\ 0 & -5 & -3 & 1 & -2 & 0 \\ 0 & 0 & 1 & 3 & -1 & -5 \end{array}\right)$$

$$R_2 \leftarrow R_2 + 3R_3: \quad \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & -6 & 3 & 10 \\ 0 & -5 & 0 & 10 & -5 & -15 \\ 0 & 0 & 1 & 3 & -1 & -5 \end{array}\right)$$

$$R_2 \leftarrow -\frac{1}{5}R_2: \quad \left(\begin{array}{ccc|ccc} 1 & 3 & 0 & -6 & 3 & 10 \\ 0 & 1 & 0 & -2 & 1 & 3 \\ 0 & 0 & 1 & 3 & -1 & -5 \end{array}\right)$$

$$R_1 \leftarrow R_1 - 3R_2: \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 & 1 & 3 \\ 0 & 0 & 1 & 3 & -1 & -5 \end{array}\right)$$

We have reduced A to the identity, and its inverse is the matrix left in the augmented segment of the augmented matrix. That is,

$$A^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ -2 & 1 & 3 \\ 3 & -1 & -5 \end{pmatrix}.$$

2.4 Exercises

1. Use Gaussian elimination to compute the inverse of

$$A = \begin{pmatrix} 1 & -1 & 2 & 0 \\ 0 & 3 & -1 & 4 \\ 2 & 0 & 1 & -3 \\ 1 & 2 & 0 & 1 \end{pmatrix}.$$

If A is not invertible, explain why.

2. Use Gaussian elimination to compute the inverse of

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 8 \\ 3 & 7 & 11 \end{pmatrix}.$$

   If A is not invertible, explain why.

3. Prove Proposition 2.1.

4. Show Proposition 2.2.

5. Compute the inverse for

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

   and write the elementary matrices corresponding to the Gaussian elimination steps used in your row reduction.

Bonus. Determine the probability of a uniformly random matrix in $M_2(\mathbb{R})$ being invertible.