Fall 2019 Math 201-003 at Drexel: blackboard notes of 2019-10-16

Darij Grinberg

December 5, 2019

1. Gaussian elimination (continued)

1.1. Invertibility (continued)

Recall:

Definition 1.1.1. Let A be an $m \times n$-matrix.

(a) A **left inverse** of A means an $n \times m$-matrix L such that $LA = I_n$.

(b) A **right inverse** of A means an $n \times m$-matrix R such that $AR = I_m$.

(c) An **inverse** (or **two-sided inverse**) of A means an $n \times m$-matrix B such that $BA = I_n$ and $AB = I_m$.

Theorem 1.1.2. Let A be a matrix. Let L be a left inverse of A. Let R be a right inverse of A. Then, $L = R$, and furthermore this matrix $L = R$ is an inverse of A.

Corollary 1.1.3. Assume that a matrix A has both a left inverse and a right inverse. Then:

(a) The matrix A has a **unique** inverse.

(b) This inverse is the **only** left inverse of A and the **only** right inverse of A.

Corollary 1.1.4. If A has an inverse, then this inverse is unique.

Definition 1.1.5. If a matrix A has an inverse, then this inverse is denoted by A^{-1}. (This notation is harmless, because the previous corollary says that this inverse is unique.)

A matrix is said to be **invertible** if it has an inverse.

Proposition 1.1.6. Let A and B be two invertible matrices. Then, their product AB is also invertible, and its inverse is $(AB)^{-1} = B^{-1}A^{-1}$.

Once again: Left inverses don't have to exist nor be unique; right inverses don't have to exist nor be unique; inverses (in the two-sided sense) don't have to exist, but at least they are unique.

Lemma 1.1.7. Let A be a matrix. Then, A has a right inverse if and only if A^T has a left inverse.

Proof. We have to prove the "if" part and the "only if" part.

"if" part: Assume that A^T has a left inverse. We must prove that A has a right inverse.

We have assumed that A^T has a left inverse; let B be such a left inverse. Thus, $BA^T = I$. (Recall that I denotes an identity matrix of any size.) But any two matrices X and Y satisfy $(XY)^T = Y^T X^T$ (provided that XY makes sense). Applying this to $X = B$ and $Y = A^T$, we find $\left(BA^T\right)^T = \underbrace{\left(A^T\right)^T}_{=A} B^T = AB^T$. Thus, $AB^T = \left(\underbrace{BA^T}_{=I}\right)^T = I^T = I$. In other words, B^T is a right inverse of A. Thus, A has a right inverse. This proves the "if" part.

"only if" part: Assume that A has a right inverse. We must prove that A^T has a left inverse.

We have assumed that A has a right inverse; let R be such a right inverse. Thus, $AR = I$. But any two matrices X and Y satisfy $(XY)^T = Y^T X^T$ (provided that XY makes sense). Applying this to $X = A$ and $Y = R$, we find $(AR)^T = R^T A^T$. Thus, $R^T A^T = \left(\underbrace{AR}_{=I}\right)^T = I^T = I$. In other words, R^T is a left inverse of A^T. Thus, A^T has a left inverse. This proves the "only if" part. □

This proof actually shows something more:

Corollary 1.1.8. Let A be a matrix. Then, a right inverse of A is the same as the transpose of a left inverse of A^T.

1.2. The Inverse Matrix Theorem

The "Inverse Matrix Theorem" is not so much a specific theorem, but a cloud of theorems, each of which has the form "a matrix is invertible if and only if (a bunch of equivalent conditions hold)". In other words, the Inverse Matrix Theorem is a set of (necessary and sufficient) criteria for a matrix to be invertible. The exact criteria depend on the author.

For us, the Inverse Matrix Theorem is the following theorem (a slightly restated version of [Strickland, Theorem 11.5]):

Theorem 1.2.1. (The Inverse Matrix Theorem.)

Let A be an $n \times n$-matrix.

Then, the following statements are equivalent (i.e., if any of them holds, then so do all the others):

- **(a)** The matrix A can be row-reduced to I_n.
- **(b)** The columns of A are linearly independent.
- **(c)** The columns of A span $\mathbb{R}^n$.
- **(d)** The columns of A form a basis of $\mathbb{R}^n$.
- **(e)** The matrix A^T can be row-reduced to I_n.
- **(f)** The columns of A^T are linearly independent.
- **(g)** The columns of A^T span $\mathbb{R}^n$.
- **(h)** The columns of A^T form a basis of $\mathbb{R}^n$.
- **(i)** The matrix A has a left inverse.
- **(j)** The matrix A has a right inverse.
- **(k)** The matrix A has an inverse (i.e., A is invertible).

Proof. The equivalence **(d)** $\Longleftrightarrow$ **(h)** follows from Proposition 1.2.6 from 2019-10-09.

The equivalence **(b)** $\Longleftrightarrow$ **(c)** $\Longleftrightarrow$ **(d)** follows from Proposition 1.2.7 from 2019-10-09. (Here we are using that our matrix A is square, so that its columns are n vectors in $\mathbb{R}^n$.)

The equivalence **(a)** $\Longleftrightarrow$ **(b)** follows from Theorem 1.1.7 from 2019-10-07. Indeed, this theorem entails that the columns of A are independent if and only if the RREF of A has a pivot in each column. But the RREF of A is a square matrix, and thus has a pivot in each column if and only if it is I_n (by Lemma 1.1.14 **(b)** from 2019-10-07).

Now, we know the equivalence **(a)** $\Longleftrightarrow$ **(b)** $\Longleftrightarrow$ **(c)** $\Longleftrightarrow$ **(d)** $\Longleftrightarrow$ **(h)**.

Applying the same reasoning to A^T instead of A, we obtain **(e)** $\Longleftrightarrow$ **(f)** $\Longleftrightarrow$ **(g)** $\Longleftrightarrow$ **(h)** $\Longleftrightarrow$ **(d)** (since $\left(A^T\right)^T = A$).

Next we shall prove the equivalence **(a)** $\Longleftrightarrow$ **(i)**. To do so, we will prove the implications **(a)** $\Longrightarrow$ **(i)** and **(i)** $\Longrightarrow$ **(a)**.

*Proof of **(a)** $\Longrightarrow$ **(i)**:* Assume that **(a)** holds. That is, the matrix A can be row-reduced to I_n. So I_n can be obtained from A by a sequence of row operations. Hence, Corollary 1.3.4 from 2019-10-09 yields that $I_n = UA$ for some matrix U that can be written as a product of elementary matrices. Consider this U. Thus, U is a left inverse of A. Thus, A has a left inverse. In other words, **(i)** holds.

*Proof of **(i)** $\Longrightarrow$ **(a)**:* Assume that **(i)** holds. That is, the matrix A has a left inverse

B. We must prove that **(a)** holds. Since we already know that **(a)** $\Longleftrightarrow$ **(b)**, it suffices to prove that **(b)** holds. In other words, we have to prove that the columns of A are independent.

Let $v_1, v_2, \ldots, v_n$ be the columns of A. Thus, $A = [v_1 \mid v_2 \mid \cdots \mid v_n]$. Hence, for each vector $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}$, we have

$$A\lambda = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n \qquad \text{(by Lemma 1.1.6 from 2019-10-07)}.$$

Thus, if

$$\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$$

is some relation between the $v_1, v_2, \ldots, v_n$, then

$$A\lambda = 0, \qquad \text{and thus}$$
$$B \underbrace{A\lambda}_{=0} = B \cdot 0 = 0, \qquad \text{hence}$$
$$0 = \underbrace{BA}_{\substack{=I_n \\ \text{(since } B \text{ is a left inverse of } A\text{)}}} \lambda = I_n \lambda = \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix},$$

so that $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$. That is, any relation between $v_1, v_2, \ldots, v_n$ is trivial. This means that $v_1, v_2, \ldots, v_n$ are independent. In other words, **(b)** holds. Since **(a)** $\Longleftrightarrow$ **(b)**, this shows that **(a)** holds.

Thus, we have proven the implications **(a)** $\Longrightarrow$ **(i)** and **(i)** $\Longrightarrow$ **(a)**. Hence, **(a)** $\Longleftrightarrow$ **(i)**.

Applying the same argument to A^T instead of A, we obtain the equivalence **(e)** $\Longleftrightarrow$ **(j)**, because we know (from the Lemma above) that A has a right inverse if and only if A^T has a left inverse.

Next, the implication **(k)** $\Longrightarrow$ **(i)** is obvious, since an inverse of A is automatically a left inverse of A.

To prove the implication **(i)** $\Longrightarrow$ **(k)**, we assume that statement **(i)** holds. Thus, A has a left inverse. But since we already know that **(i)** $\Longleftrightarrow$ **(j)** (because **(i)** $\Longleftrightarrow$ **(a)** $\Longleftrightarrow$ **(e)** $\Longleftrightarrow$ **(j)**), we know that statement **(j)** holds as well. In other words, A has a right inverse. So we know that A has both a left inverse and a right inverse. Thus, by Corollary 1.4.4 **(a)** from 2019-10-09, we conclude that A has an inverse. Thus, **(k)** holds.

So we have proven the implications **(k)** $\Longrightarrow$ **(i)** and **(i)** $\Longrightarrow$ **(k)**. Thus, **(i)** $\Longleftrightarrow$ **(k)**.

Combining the equivalences we have shown, we can now conclude that all our statements are equivalent. $\square$

Example 1.2.2. Is the matrix $A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ invertible?

According to the Inverse Matrix Theorem (specifically, the equivalence **(a)** $\Longleftrightarrow$ **(k)**), this boils down to checking whether A can be row-reduced to $I_n = I_2$. Let us check this:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \overset{\text{add } -3\cdot\text{row } 1 \text{ to row } 2}{\longrightarrow} \begin{pmatrix} \boxed{1} & 2 \\ 0 & \boxed{-2} \end{pmatrix} \overset{\text{scale row 2 by } -1/2}{\longrightarrow} \begin{pmatrix} \boxed{1} & 2 \\ 0 & \boxed{1} \end{pmatrix}$$
$$\overset{\text{add } -2\cdot\text{row } 2 \text{ to row } 1}{\longrightarrow} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2.$$

This shows that A is invertible.

What is A^{-1} ? Well, let us follow our proof of **(a)** $\Longrightarrow$ **(i)** above, and actually compute the left inverse of A (which, as we know, will be an inverse of A). So we look at our above way of row-reducing A to I_2, and rewrite it in terms of multiplication by elementary matrices (using Proposition 1.3.3 from 2019-10-09):

$$\begin{pmatrix} \boxed{1} & 2 \\ 0 & \boxed{-2} \end{pmatrix} = U_1 A \qquad \text{for } U_1 = E_{2,1}(-3);$$
$$\begin{pmatrix} \boxed{1} & 2 \\ 0 & \boxed{1} \end{pmatrix} = U_2 \begin{pmatrix} \boxed{1} & 2 \\ 0 & \boxed{-2} \end{pmatrix} \qquad \text{for } U_2 = D_2(-1/2);$$
$$I_2 = U_3 \begin{pmatrix} \boxed{1} & 2 \\ 0 & \boxed{1} \end{pmatrix} \qquad \text{for } U_3 = E_{1,2}(-2).$$

Thus,
$$I_2 = U_3 U_2 U_1 A.$$

Hence, $U_3 U_2 U_1$ is a left inverse of A, and therefore must be the inverse of A (since A is invertible). Thus,

$$\begin{aligned} A^{-1} &= U_3 U_2 U_1 = E_{1,2}(-2)\, D_2(-1/2)\, E_{2,1}(-3) \\ &= \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1/2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}. \end{aligned}$$

Example 1.2.3. Now, let $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. Is A invertible?

Again, use statement **(a)** from the Inverse Matrix Theorem:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \overset{\text{add } -3\cdot\text{row } 1 \text{ to row } 2}{\longrightarrow} \begin{pmatrix} \boxed{1} & 2 \\ 0 & 0 \end{pmatrix}.$$

This is a RREF, but not I_n. Thus, statement **(a)** does **not** hold here. Hence, statement **(k)** does not hold either. That is, A is not invertible.

Proposition 1.2.4. Any elementary matrix is invertible.

Proof. (This is [Strickland, Example 11.8].)
We need to show that $D_p(\lambda)$, $E_{p,q}(\mu)$ and $F_{p,q}$ are invertible.

- For $D_p(\lambda)$, we claim that $D_p(\lambda) \cdot D_p(\lambda^{-1}) = I_n$. Indeed:
  - Proposition 1.3.3 **(a)** from 2019-10-09 shows that $D_p(\lambda^{-1}) = D_p(\lambda^{-1}) \cdot I_n$ is obtained from I_n by scaling the p-th row by λ^{-1}. (Of course, this is also obvious from the definition of $D_p(\lambda^{-1})$.)
  - Proposition 1.3.3 **(a)** from 2019-10-09 shows that $D_p(\lambda) \cdot D_p(\lambda^{-1})$ is obtained from $D_p(\lambda^{-1})$ by scaling the p-th row by λ.

  Thus, $D_p(\lambda) \cdot D_p(\lambda^{-1})$ is obtained from I_n by first scaling the p-th row by λ^{-1} and then scaling it by λ again. Clearly, these two scaling operations undo each other, so after doing both of them we just get I_n back. So $D_p(\lambda) \cdot D_p(\lambda^{-1}) = I_n$.

  Similarly, $D_p(\lambda^{-1}) \cdot D_p(\lambda) = I_n$. Thus, $D_p(\lambda^{-1})$ is an inverse of $D_p(\lambda)$.

- For $E_{p,q}(\mu)$, we claim that $E_{p,q}(\mu) \cdot E_{p,q}(-\mu) = I_n$. Indeed:
  - Proposition 1.3.3 **(b)** from 2019-10-09 shows that $E_{p,q}(-\mu) = E_{p,q}(-\mu) \cdot I_n$ is obtained from I_n by adding $-\mu$ times the q-th row to the p-th row.
  - Proposition 1.3.3 **(b)** from 2019-10-09 shows that $E_{p,q}(\mu) \cdot E_{p,q}(-\mu)$ is obtained from $E_{p,q}(-\mu)$ by adding μ times the q-th row to the p-th row.

  Thus, $E_{p,q}(\mu) \cdot E_{p,q}(-\mu)$ is obtained from I_n by first adding $-\mu$ times the q-th row to the p-th row, and then adding μ times the q-th row to the p-th row. Clearly, these two adding operations undo each other, so after doing both of them we just get I_n back. So $E_{p,q}(\mu) \cdot E_{p,q}(-\mu) = I_n$.

  Similarly, $E_{p,q}(-\mu) \cdot E_{p,q}(\mu) = I_n$. Thus, $E_{p,q}(-\mu)$ is an inverse of $E_{p,q}(\mu)$.

- For $F_{p,q}$, we claim that $F_{p,q} \cdot F_{p,q} = I_n$. The proof is similar. Thus, $F_{p,q}$ is its own inverse.

□

So we know how to find inverses of square matrices. What about rectangular matrices?

Theorem 1.2.5. (The Left Inverse Matrix Theorem)
Let A be an $n \times m$-matrix. Then, the following statements are equivalent:

- **(a)** The RREF of A has a pivot in each column.
- **(b)** The columns of A are linearly independent.
- **(g)** The columns of A^T span $\mathbb{R}^m$.
- **(i)** The matrix A has a left inverse.

Proof. **(a)** $\Longleftrightarrow$ **(b)** is proved as in the Inverse Matrix Theorem.
(b) $\Longleftrightarrow$ **(g)** is Proposition 1.1.3 **(a)** from 2019-10-09.
(i) $\Longrightarrow$ **(a)** is proved as in the Inverse Matrix Theorem.
Let us now prove **(a)** $\Longrightarrow$ **(i)**. So we assume that **(a)** holds. That is, the RREF of A has a pivot in each column. Let us denote this RREF by B. Then, B is a RREF matrix with a pivot in each column. Hence, Lemma 1.1.14 **(c)** from 2019-10-07 shows that $n \geq m$ and

$$B = \begin{bmatrix} I_m \\ 0_{(n-m)\times m} \end{bmatrix}.$$

Let C be the $m \times n$-matrix whose leftmost m columns form the identity matrix I_m and its remaining $n - m$ columns are just zero. (Note that $C = B^T$, but we won't need this.) It is easy to see that $CB = I_m$.
But B can be obtained from A by a sequence of row operations (since B is the RREF of A). Hence, Corollary 1.3.4 from 2019-10-09 yields that $B = UA$ for some matrix U that can be written as a product of elementary matrices. Consider this U.
Now, $(CU)\,A = C\underbrace{UA}_{=B} = CB = I_m$. Hence, CU is a left inverse of A. Thus, A has a left inverse. In other words, **(i)** holds. This proves the implication **(a)** $\Longrightarrow$ **(i)**.
Combining this with **(i)** $\Longrightarrow$ **(a)**, we obtain the equivalence **(a)** $\Longleftrightarrow$ **(i)**. By combining what we have proved, we get the full set of equivalences **(a)** $\Longleftrightarrow$ **(b)** $\Longleftrightarrow$ **(g)** $\Longleftrightarrow$ **(i)**. $\square$

Theorem 1.2.6. (The Right Inverse Matrix Theorem)
Let A be an $n \times m$-matrix. Then, the following statements are equivalent:

- **(e)** The RREF of A^T has a pivot in each column.
- **(f)** The columns of A^T are linearly independent.
- **(c)** The columns of A span $\mathbb{R}^n$.
- **(j)** The matrix A has a right inverse.

Proof. Apply the Left Inverse Matrix Theorem to A^T instead of A. Again, the trick is that A has a right inverse if and only if A^T has a left inverse (by the first Lemma we proved today). □

Corollary 1.2.7. Let A be an $n \times m$-matrix.
(a) If A has a left inverse, then $n \geq m$ (that is, A is square or tall).
(b) If A has a right inverse, then $n \leq m$ (that is, A is square or wide).
(c) If A has an inverse, then $n = m$ (that is, A is square).

Proof. **(a)** Assume that A has a left inverse. Then, the Left Inverse Matrix Theorem says that the columns of A are linearly independent. Thus, they are m linearly independent vectors in $\mathbb{R}^n$. If $m > n$, then Corollary 1.2.9 from 2019-10-09 would say that this is impossible. Thus, we must have $m \leq n$, that is, $n \geq m$.

(b) This follows from part **(a)**, applied to A^T instead of A.

(c) Follows by combining **(a)** with **(b)** (since an inverse is both a left inverse and a right inverse). □

1.3. Finding inverses

Is there a better algorithm for inverting a matrix (i.e., finding its inverse) than what we did in the examples above?

Yes. We will use the following notation:

Definition 1.3.1. Let B be an $n \times m$-matrix, and let C be an $n \times p$-matrix. Then, $[B \mid C]$ denotes the $n \times (m+p)$-matrix obtained by gluing C to the right edge of B.

Example 1.3.2. If $B = \begin{pmatrix} u & v \\ w & x \end{pmatrix}$ and $C = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$, then

$$[B \mid C] = \begin{pmatrix} u & v & a & b & c \\ w & x & d & e & f \end{pmatrix}.$$

More generally, we can define the notation $[B_1 \mid B_2 \mid \cdots \mid B_m]$ for any (finite) sequence of matrices $B_1, B_2, \ldots, B_m$ with the same number of rows. This generalizes the notation $[v_1 \mid v_2 \mid \cdots \mid v_m]$ for m column vectors $v_1, v_2, \ldots, v_m$.

Lemma 1.3.3. Let A be a $q \times n$-matrix. Let B be an $n \times m$-matrix, and let C be an $n \times p$-matrix. Then,

$$A\,[B \mid C] = [AB \mid AC].$$

Proof. The definition of $[B \mid C]$ yields

$$[B \mid C] = \left[\operatorname{col}_1 B \mid \operatorname{col}_2 B \mid \cdots \mid \operatorname{col}_m B \mid \operatorname{col}_1 C \mid \operatorname{col}_2 C \mid \cdots \mid \operatorname{col}_p C\right]$$

(where we use our old notation $\operatorname{col}_k D$ for the k-th column of a matrix D).

Recall (from Proposition 2.6.2 **(c)** from 2019-09-23) that

$$\operatorname{col}_j(AB) = A \cdot \operatorname{col}_j B \qquad \text{for each } j \in \{1, 2, \ldots, m\};$$

thus,

$$AB = [A \cdot \operatorname{col}_1 B \mid A \cdot \operatorname{col}_2 B \mid \cdots \mid A \cdot \operatorname{col}_m B].$$

Similarly,

$$AC = [A \cdot \operatorname{col}_1 C \mid A \cdot \operatorname{col}_2 C \mid \cdots \mid A \cdot \operatorname{col}_p C]$$

and

$$\begin{aligned}
&A[B \mid C] \\
&= [A \cdot \operatorname{col}_1 [B \mid C] \mid A \cdot \operatorname{col}_2 [B \mid C] \mid \cdots \mid A \cdot \operatorname{col}_{m+p} [B \mid C]] \\
&= [A \cdot \operatorname{col}_1 B \mid A \cdot \operatorname{col}_2 B \mid \cdots \mid A \cdot \operatorname{col}_m B \mid A \cdot \operatorname{col}_1 C \mid A \cdot \operatorname{col}_2 C \mid \cdots \mid A \cdot \operatorname{col}_p C] \\
&\qquad \left(\text{since } [B \mid C] = [\operatorname{col}_1 B \mid \operatorname{col}_2 B \mid \cdots \mid \operatorname{col}_m B \mid \operatorname{col}_1 C \mid \operatorname{col}_2 C \mid \cdots \mid \operatorname{col}_p C]\right) \\
&= [AB \mid AC].
\end{aligned}$$

□

Example 1.3.4. If $A = \begin{pmatrix} 1 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} u & v \\ w & x \end{pmatrix}$ and $C = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$, then

$$AB = \begin{pmatrix} u + 2w & v + 2x \end{pmatrix} \qquad \text{and} \qquad AC = \begin{pmatrix} a + 2d & b + 2e & c + 2f \end{pmatrix},$$

thus

$$[AB \mid AC] = \begin{pmatrix} u + 2w & v + 2x & a + 2d & b + 2e & c + 2f \end{pmatrix}.$$

On the other hand,

$$[B \mid C] = \begin{pmatrix} u & v & a & b & c \\ w & x & d & e & f \end{pmatrix}, \qquad \text{so that}$$

$$A[B \mid C] = \begin{pmatrix} u + 2w & v + 2x & a + 2d & b + 2e & c + 2f \end{pmatrix}.$$

The following method for inverting a matrix is [Strickland, Method 11.11]:

Theorem 1.3.5. Let A be an $n \times n$-matrix. Form the $n \times (2n)$-matrix $[A \mid I_n]$. Apply row operations to bring the latter matrix into RREF; let this RREF be $[T \mid B]$, where T is its "left half" (i.e., the first n columns) and B is its "right half" (i.e., the last n columns).

(a) If $T = I_n$, then A is invertible and its inverse is $A^{-1} = B$.

(b) If $T \neq I_n$, then A is not invertible.

Example 1.3.6. Let $n = 2$ and $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Let us bring $[A \mid I_n]$ to RREF:

$$[A \mid I_n] = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{pmatrix} \overset{\text{add } -3\cdot\text{row 1 to row 2}}{\longrightarrow} \begin{pmatrix} \boxed{1} & 2 & 1 & 0 \\ 0 & \boxed{-2} & -3 & 1 \end{pmatrix}$$

$$\overset{\text{scale row 2 by } -1/2}{\longrightarrow} \begin{pmatrix} \boxed{1} & 2 & 1 & 0 \\ 0 & \boxed{1} & 3/2 & -1/2 \end{pmatrix}$$

$$\overset{\text{add } -2\cdot\text{row 2 to row 2}}{\longrightarrow} \begin{pmatrix} \boxed{1} & 0 & -2 & 1 \\ 0 & \boxed{1} & 3/2 & -1/2 \end{pmatrix}$$

$$= [T \mid B] \qquad \text{for } T = \begin{pmatrix} \boxed{1} & 0 \\ 0 & \boxed{1} \end{pmatrix} \text{ and } B = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}.$$

Since $T = I_2$, we thus conclude that A is invertible, and its inverse is $A^{-1} = B = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$.

Proof of Theorem. We have obtained $[T \mid B]$ from $[A \mid I_n]$ by row operations. Thus, Corollary 1.3.4 from 2019-10-09 shows that

$$[T \mid B] = U [A \mid I_n]$$

for some matrix U that can be written as a product of elementary matrices. Consider this U. Now,

$$[T \mid B] = U [A \mid I_n] = [UA \mid UI_n] \qquad \text{(by the last Lemma)}.$$

In other words, $T = UA$ and $B = UI_n$. Hence, $B = UI_n = U$.

(a) If $T = I_n$, then we thus have $I_n = T = \underbrace{U}_{=B} A = BA$, which shows that B is a left inverse of A, and thus A has an inverse (by the Inverse Matrix Theorem), which must therefore be B (since $I_n = BA$).

(b) Assume that $T \neq I_n$. We must show that A is not invertible.

We know that $[T \mid B]$ can be obtained from $[A \mid I_n]$ by row operations.

Thus, T can be obtained from A by the same row operations. Moreover, since $[T \mid B]$ is in RREF, T must also be in RREF, because T is the left half of $[T \mid B]$. Thus, T is the RREF of A. Since $T \neq I_n$, we conclude that A cannot be row-reduced to I_n. Thus, statement **(a)** of the Inverse Matrix Theorem is false. Hence, statement **(k)** is false. In other words, A is not invertible. □

Example 1.3.7. Let

$$A = \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}.$$

Then, A is easily seen to be invertible; what is its inverse?

$$\begin{aligned}
[A \mid I_n] &= \begin{pmatrix} 1 & a & b & 1 & 0 & 0 \\ 0 & 1 & c & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \\
&\longrightarrow \begin{pmatrix} 1 & a & b & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & -c \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \\
&\longrightarrow \begin{pmatrix} 1 & a & 0 & 1 & 0 & -b \\ 0 & 1 & 0 & 0 & 1 & -c \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} \\
&\longrightarrow \begin{pmatrix} 1 & 0 & 0 & 1 & -a & -b - a(-c) \\ 0 & 1 & 0 & 0 & 1 & -c \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.
\end{aligned}$$

Thus, A is invertible, and its inverse is

$$A^{-1} = \begin{pmatrix} 1 & -a & -b - a(-c) \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -a & ac - b \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix}.$$

In the same way, you can see that **any** upper-triangular matrix with nonzero entries on the diagonal is invertible, and its inverse is again an upper-triangular matrix with nonzero entries on the diagonal.

References

[Strickland] Neil Strickland, *MAS201 Linear Mathematics for Applications*, lecture notes, 28 September 2013.
`http://neil-strickland.staff.shef.ac.uk/courses/MAS201/`