Recall that any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ can be represented in terms of the standard basis by an $n \times n$ matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}.$$ 

In other words, the linear transformation can be represented in terms of the vector

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad \text{as} \quad T(x) = Ax.$$

We can think of the linear transformation as taking $x$ to

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

If the linear transformation is one to one and onto, it possesses an inverse map $T^{-1}$, which is also a linear transformation, and is represented by a matrix

$$B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix},$$

called the inverse of $A$. How do we find this inverse matrix?

The idea is very simple. We think of the linear transformation from $\mathbb{R}^n$ to itself in terms of equations

$$a_{11}x_1 + \ldots + a_{1n}x_n = y_1,$$

$$a_{21}x_1 + \ldots + a_{2n}x_n = y_2,$$

$$\ldots$$

$$a_{n1}x_1 + \ldots + a_{nn}x_n = y_n,$$
and solve for the variables \((x_1, \ldots, x_n)\) in terms of \((y_1, \ldots, y_n)\). This can be done if the linear transformation is invertible.

As might be expected, the procedure for inverting a linear transformation consists of applying the elementary row operations. For example, to invert the linear transformation

\[
\begin{align*}
  x_1 + x_2 &= y_1, \\
  2x_1 + x_2 &= y_2,
\end{align*}
\]

one subtracts twice the first equation from the second,

\[
\begin{align*}
  x_1 + x_2 &= y_1, \\
  -x_2 &= -2y_1 + y_2,
\end{align*}
\]

and then adds the second from the first,

\[
\begin{align*}
  x_1 &= -y_1 + y_2, \\
  -x_2 &= -2y_1 + y_2.
\end{align*}
\]

Finally, one multiplies the second equation by \(-1\), obtaining the inverse transformation

\[
\begin{align*}
  x_1 &= -y_1 + y_2, \\
  x_2 &= 2y_1 - y_2.
\end{align*}
\]

In matrix notation, the equations

\[
\begin{pmatrix}
  1 & 1 \\
  2 & 1
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2
\end{pmatrix}
= 
\begin{pmatrix}
  y_1 \\
  y_2
\end{pmatrix}
\text{ and }
\begin{pmatrix}
  x_1 \\
  x_2
\end{pmatrix}
= 
\begin{pmatrix}
  -1 & 1 \\
  2 & -1
\end{pmatrix}
\begin{pmatrix}
  y_1 \\
  y_2
\end{pmatrix}
\]

are equivalent. The reader can verify that the matrices

\[
A = \begin{pmatrix}
  1 & 1 \\
  2 & 1
\end{pmatrix}
\text{ and } B = \begin{pmatrix}
  -1 & 1 \\
  2 & -1
\end{pmatrix}
\]

do indeed satisfy the equations

\[
AB = BA = I,
\]

where \(I\) is the \(n \times n\) identity matrix.

The method can be streamlined by writing only the matrices, and not the variables. One applies the elementary row operations to the \(n \times (2n)\) matrix \((A|I)\) obtained by placing \(A\) and the \(n \times n\) identity matrix side by side. The goal is to put the resulting matrix in row-reduced form. If \(A\) is invertible, we will be able to find a row-equivalent matrix of the form \((I|B)\), and \(B\) will be the inverse of \(A\).

Let us illustrate the procedure with the matrix

\[
A = \begin{pmatrix}
  1 & 1 \\
  2 & 1
\end{pmatrix}.
\]

We apply the elementary row operations to the matrix

\[
(A|I) = \begin{pmatrix}
  1 & 1 & | & 1 & 0 \\
  2 & 1 & | & 0 & 1
\end{pmatrix},
\]

2
obtaining
\[
\begin{pmatrix}
1 & 1 & & 1 & 0 \\
0 & -1 & & -2 & 1
\end{pmatrix} \mapsto \begin{pmatrix}
1 & 0 & & -1 & 1 \\
0 & -1 & & -2 & 1
\end{pmatrix} \mapsto \begin{pmatrix}
1 & 0 & & -1 & 1 \\
0 & 1 & & 2 & -1
\end{pmatrix}.
\]
we have replaced the matrix \((A|I)\) by a matrix of the form \((I|B)\). In this case
\(A\) is invertible, and its inverse is
\[
B = \begin{pmatrix}
-1 & 1 \\
2 & -1
\end{pmatrix}.
\]
We will often denote the inverse of \(A\) by \(A^{-1}\).

We have illustrated the procedure for finding inverses with \(2 \times 2\) matrices. Exactly the same procedure works for \(n \times n\) matrices, but of course the arithmetic is often more complicated.

There is a shortcut that works for \(2 \times 2\) matrices. If
\[
A = \begin{pmatrix}
a & b \\
c & d
\end{pmatrix}, \quad \text{then} \quad A^{-1} = \frac{1}{ad - bc} \begin{pmatrix}
d & -b \\
-c & a
\end{pmatrix},
\]
You can check for yourself that this formula works.

Sometimes we can find the inverse with very little work. For example, if \(\theta\) is an angle expressed in radians, the linear transformation \(T : \mathbb{R}^n \to \mathbb{R}^n\) represented by the matrix
\[
A = \begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
\]
is a counterclockwise rotation through the angle \(\theta\). The linear transformation which undoes this is just a clockwise rotation through the angle \(\theta\), which corresponds to replacing \(\theta\) by \(-\theta\). So
\[
A^{-1} = \begin{pmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{pmatrix}
\]
Of course, you could verify this with (2).