We begin with a proof of the fact that

$$(AB)^T = B^T A^T$$

Case I. B is a column matrix

$$(AB)^T = \left(\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right)^T$$

$$= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}^T$$

$$= [a_{11}x_1 + \cdots + a_{1n}x_n, a_{21}x_1 + \cdots + a_{2n}x_n, \cdots; a_{m1}x_1 + \cdots + a_{mn}x_n]$$

$$= [x_1, x_2, \cdots, x_n] \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} = B^T A^T$$

Case II. B has columns $B_1, B_2, \ldots, B_p$

$$(AB)^T = (A[B_1, B_2, \ldots, B_p])^T = [AB_1, AB_2, \ldots, AB_p]^T$$

$$= \begin{bmatrix} (AB_1)^T \\ (AB_2)^T \\ \vdots \\ (AB_p)^T \end{bmatrix} = \begin{bmatrix} B_1^T A^T \\ B_2^T A^T \\ \vdots \\ B_p^T A^T \end{bmatrix} = \begin{bmatrix} B_1^T \\ B_2^T \\ \vdots \\ B_p^T \end{bmatrix} A^T = B^T A^T$$

Q.E.D.

It follows that $AB = I \Rightarrow B^T A^T = I$ and $BA = I \Rightarrow A^T B^T = I$. Hence

A invertible $\Rightarrow A^T$ invertible and $(A^T)^{-1} = (A^{-1})^T$

If $BA = I$ it follows that $AX = 0 \Rightarrow X = 0$ on multiplying both sides of $AX = 0$ by B. Hence, in this case, the number of columns of A cannot exceed the number of rows since a homogeneous system of linear equations with more unknowns than equations always has a non-zero solution. Thus a matrix having more columns than rows cannot be invertible. A matrix having more rows than columns also cannot be invertible; otherwise A^T would be invertible and A^T would have more columns than rows.

A matrix is said to be square if it has the same number of rows and columns. A non-square matrix cannot be invertible.

We now prove the following result

Theorem. A matrix A is invertible $\Longleftrightarrow$ A is row equivalent to the identity matrix i.e A can be row reduced to the identity matrix

Proof. If A is row equivalent to the identity matrix I then there is a matrix B such that $AX = Y \Longleftrightarrow X = BY$. We claim that B is the inverse of A. Indeed, for any Y we have

$$Y = AX = A(BY) = (AB)Y$$

If I_i is the ith column of I, we get

$I_i = (AB)I_i =$ i-th column of AB

Hence $AB = I$. Similarly, for any X,

$X = BY = B(AX) = (BA)X \Rightarrow BA = I$.

Conversely, if B is a matrix with $BA = I$ then $AX = 0 \Rightarrow X = 0$ which shows that A must be row equivalent to I.

Q.E.D

It follows from the proof that, for a square matrix, $BA = I \Rightarrow AB = I$ and hence A invertible.

Exercise. Show that $BA = I \Rightarrow AB = I$ and hence that A is invertible.

We now give a different proof of the above theorem using elementary matrices.

Definition. An elementary matrix is a matrix that can be obtained from the identity matrix by a single elementary operation.

Example. The elementary 2×2 matrices are

$$\begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}, \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & c \end{bmatrix} \text{ with } c \neq 0$$

Theorem. If ξ is an elementary operation and $A \xrightarrow{\xi} B$, $I \xrightarrow{\xi} E$ then $B = EA$

Proof. Case I: ξ is the elem. op. $R_i \to R_i + aR_j$. Then, if I_i is the i-th row of I, we have

$$E = \begin{bmatrix} I_1 \\ \vdots \\ I_i + aI_j \\ \vdots \\ I_n \end{bmatrix}$$

and the i-th row of EA is equal to $(I_i + aI_j)A = I_iA + aI_jA =$ i-th row of A + $a \times$ j-th row of A.

If $k \neq i$ then the k-th row of EA is I_kA which is the k-th row of A. Hence $EA = B$.

The cases where ξ is $R_i \leftrightarrow R_j$ or $R_i \to cR_j$ is left to the reader.

Example. $\begin{bmatrix}1&2\\2&6\end{bmatrix} \xrightarrow{R_2-2R_1} \begin{bmatrix}1&2\\0&2\end{bmatrix}$

Here $\mathcal{E}$ is the elem. op. $R_2 \to R_2 - 2R_1$. Since $\begin{bmatrix}1&0\\0&1\end{bmatrix} \xrightarrow{R_2-2R_1} \begin{bmatrix}1&0\\-2&1\end{bmatrix}$ the corresponding elementary matrix is $\begin{bmatrix}1&0\\-2&1\end{bmatrix}$ and

$$\begin{bmatrix}1&0\\-2&1\end{bmatrix}\begin{bmatrix}1&2\\2&6\end{bmatrix} = \begin{bmatrix}1&2\\0&2\end{bmatrix}$$

We now show that elementary matrices are invertible with inverses that are also elementary matrices. If $\mathcal{E}$ is an elementary operation let $\mathcal{F}$ be the inverse operation. If $\mathcal{E}$ is $R_i \to R_i + aR_j$ then $\mathcal{F}$ is $R_i \to R_i - aR_j$; if $\mathcal{E}$ is $R_i \leftrightarrow R_j$ then $\mathcal{F} = \mathcal{E}$; if $\mathcal{E}$ is $R_i \to cR_i$ then $\mathcal{F}$ is $R_i \to \frac{1}{c}R_i$.

If $I \xrightarrow{\mathcal{E}} E$ and $I \xrightarrow{\mathcal{F}} F$ then $I \xrightarrow{\mathcal{E}} E \xrightarrow{\mathcal{F}} FE = I$ and $I \xrightarrow{\mathcal{F}} F \xrightarrow{\mathcal{E}} EF = I \Rightarrow F = E^{-1}$.

Example. $\begin{bmatrix}1&0\\a&1\end{bmatrix}^{-1} = \begin{bmatrix}1&0\\-a&1\end{bmatrix}$, $\begin{bmatrix}1&a\\0&1\end{bmatrix}^{-1} = \begin{bmatrix}1&-a\\0&1\end{bmatrix}$

$\begin{bmatrix}0&1\\1&0\end{bmatrix}^{-1} = \begin{bmatrix}0&1\\1&0\end{bmatrix}$, $\begin{bmatrix}c&0\\0&1\end{bmatrix}^{-1} = \begin{bmatrix}1/c&0\\0&1\end{bmatrix}$, $\begin{bmatrix}1&0\\0&c\end{bmatrix}^{-1} = \begin{bmatrix}1&0\\0&1/c\end{bmatrix}$

if $c \neq 0$.