Solution (#686) It is important firstly to note that after some swapping of rows we will only be able to add multiples of a row to lower rows (using the EROs $A_{ij}(\lambda)$ where $i<j$).

We will prove this by induction on the number of rows of a matrix M.

Any $1\times n$ matrix M is already in row echelon form.

Say now that $k\times n$ matrices may be reduced to row echelon form as desired when $k<m$ and let M be an $m\times n$ matrix. Without loss of generality we may assume that the first column of M is non-zero. (Otherwise we can just focus on the submatrix made up of the first non-zero column of M and everything to its right.) By swapping the first row with a lower (ith) row if necessary, we can arrange that $[M_{11}]\neq 0$. Then by adding appropriate multiples of the first row to lower rows we can clear out the first column of M.

That is

$$A_{12}A_{13}\cdots A_{1m}S_{1i}M=\begin{pmatrix} 1 & \mathbf{r} \\ \mathbf{0} & \tilde{M} \end{pmatrix},$$

for some $(m-1)\times(n-1)$ matrix $\tilde{M}$, and where the ith row of M was $(1,\mathbf{r})$. (For notational ease we have suppressed the multiples λ involved in the A_{1k}s.)

By our inductive hypothesis there is now a sequence of EROs $\tilde{S}_{a_ib_i}$ and of the form $\tilde{A}_{c_id_i}(\lambda_i)$ where $c_i<d_i$ such that

$$\tilde{A}_{c_1d_1}\cdots\tilde{A}_{c_rd_r}\tilde{S}_{a_1b_1}\cdots\tilde{S}_{a_sb_s}\tilde{M}$$

is in row echelon form. Note that $\tilde{S}_{a_ib_i}$ and $\tilde{A}_{c_id_i}$ are represented by $(m-1)\times(m-1)$ elementary matrices and that

$$\operatorname{diag}(1,\tilde{A}_{cd})=A_{(c+1)(d+1)},\qquad \operatorname{diag}(1,\tilde{S}_{ab})=S_{(a+1)(b+1)},$$

are also elementary $m\times m$ matrices. Thus we have

$$A_{(c_1+1)(d_1+1)}\cdots A_{(c_r+1)(d_r+1)}S_{(a_1+1)(b_1+1)}\cdots S_{(a_s+1)(b_s+1)}A_{12}A_{13}\cdots A_{1m}S_{1i}M$$

is in row echelon form.

We are almost done, our only problem being that the EROs applied to put M into row echelon form are not quite in the required order. However we showed in #685 for distinct i,j,k,l that

$$S_{ij}A_{kl}(\lambda)=A_{kl}(\lambda)S_{ij};\qquad S_{ij}A_{ik}(\lambda)=A_{jk}(\lambda)\,S_{ij}.$$

So we have either

$$S_{(a_s+1)(b_s+1)}A_{12}=A_{12}S_{(a_s+1)(b_s+1)}$$

or, when a_s+1 or b_s+1 equals 2, something of the form

$$S_{(a_s+1)(b_s+1)}A_{12}=A_{1(a_s+1)}S_{(a_s+1)(b_s+1)}.$$

Importantly, for both cases, the ERO A_{12} or $A_{1(a_s+1)}$ is one of the form A_{ij} where $i<j$. As none of the values $a_1+1,b_1+1,\ldots,a_s+1,b_s+1$ equal 1 we can take each of the EROs $S_{(a_1+1)(b_1+1)},\ldots,S_{(a_s+1)(b_s+1)}$ past the EROs $A_{12},A_{13},\cdots,A_{1m}$ until we have a sequence of EROs of the desired form.

The result then follows by induction.

Remark: Consequently, then, it is possible to put A into row echelon form using a sequence of EROs of the form S_{ij} followed by a sequence of EROs of the form $A_{ij}(\lambda)$ where $i<j$. The combined effect of the S_{ij} EROs is a 'permutation matrix' P. Further the elementary matrix representing $A_{ij}(\lambda)$ where $i<j$ is lower triangular with diagonal entries equalling 1, so this is also true of their product $\tilde{L}$. As the square matrix $\tilde{L}PA=U$ is in row echelon form then U is upper triangular. Thus we have $\tilde{L}PA=U$. If we write $L=\tilde{L}^{-1}$, then L is again lower triangular with diagonal entries equalling 1 and we have $PA=LU$ as required. This exercise therefore contains most of the details of the proof of Theorem 279.