Summer Programme 2021

Solutions to exercises from Michael Artin's *Algebra*

Satvik Saha
19MS154

Indian Institute of Science Education and Research, Kolkata,
Mohanpur, West Bengal, 741246, India.

Chapter 1

Matrix Operations

1.4 Permutation Matrices

Exercise 1. Consider the permutation p defined by $1 \rightsquigarrow 3$, $2 \rightsquigarrow 1$, $3 \rightsquigarrow 4$, $4 \rightsquigarrow 2$.

(a) Find the associated permutation matrix P.

(b) Write p as a product of transpositions and evaluate the corresponding matrix product.

(c) Compute the sign of p.

Solution.

(a) The column P_i must be the standard basis vector $\boldsymbol{e}_{p(i)}$, so

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

(b) Check that $p = (1,3,4,2) = (1,2)(1,4)(1,3)$. This product is given by

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = P.$$

(c) Since p is the product of an odd number of transpositions, its sign is -1. This is verified by calculating the determinant

$$\det \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = -\det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = -\det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = -1.$$

Exercise 2. Prove that every permutation matrix is a product of transpositions.

Solution. Note that this is equivalent to stating that any ordered list can be sorted using transpositions.

The statement is trivially true for all 2×2 permutation matrices,

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

the first being the identity and the second being a transposition itself. Suppose that any $n \times n$ permutation matrix is the product of transpositions. Use the fact that for square matrices A and B,

$$\begin{bmatrix} A & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} B & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} AB & 0 \\ 0 & 1 \end{bmatrix},$$

which means that if a permutation matrix $P = E_1 E_2 \dots E_k$, then

$$\begin{bmatrix} P & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} E_1 & 0 \\ 0 & 1 \end{bmatrix} \cdots \begin{bmatrix} E_k & 0 \\ 0 & 1 \end{bmatrix}.$$

Now let Q be an arbitrary $(n+1) \times (n+1)$ permutation matrix. Let j be the index of the row of Q which is precisely $(0 \dots 0\, 1)$, and let E be the transposition matrix which interchanges the rows $j \leftrightarrow n+1$. Then,

$$EQ = \begin{bmatrix} Q' & 0 \\ 0 & 1 \end{bmatrix},$$

where Q' is an $n \times n$ permutation matrix. This is because Q' has exactly one 1 in each row and column, the remaining elements being 0. Multiply both sides by E, and use the fact that $E^2 = \mathbb{I}$. Now, Q' is a product of transpositions $E_1 \dots E_k$, so we finally have

$$Q = E \begin{bmatrix} Q' & 0 \\ 0 & 1 \end{bmatrix} = E \begin{bmatrix} E_1 & 0 \\ 0 & 1 \end{bmatrix} \cdots \begin{bmatrix} E_k & 0 \\ 0 & 1 \end{bmatrix}.$$

Exercise 3. Prove that every matrix with a single 1 in each row and a single 1 in each column, the other entries being zero, is a permutation matrix.

Solution. Note that each column of such a matrix P must be a distinct standard basis vector $\boldsymbol{e}_k$, and we claim that this matrix represents the permutation p defined as $p(j) = k$, where $P_j = \boldsymbol{e}_k$ is the j^{th} column of P. Now, p is a bijection because every column j has one and exactly one 1 in the k^{th} row. This justifies that p is indeed a permutation. When P acts on a column vector $\boldsymbol{x}$, we have

$$P\boldsymbol{x} = P_1 x_1 + P_2 x_2 + \cdots + P_n x_n = \boldsymbol{e}_{p(1)} x_1 + \boldsymbol{e}_{p(2)} x_2 + \cdots + \boldsymbol{e}_{p(n)} x_n.$$

This means that

$$P \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_{p^{-1}(1)} \\ x_{p^{-1}(2)} \\ \vdots \\ x_{p^{-1}(n)} \end{bmatrix}.$$

Exercise 4. Let p be a permutation. Prove that $\operatorname{sign} p = \operatorname{sign} p^{-1}$.

Solution. This follows directly from the fact that $\det P^{-1} = 1/\det P$, and that $\det P = \pm 1$ so $\det P^{-1} = \det P$.

Exercise 5. Prove that the transpose of a permutation matrix P is its inverse.

Solution. Recall that $\det P = \pm 1$, so P is invertible. Write the permutation matrix P in terms of its columns,

$$P = \begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{e}_{p(1)} & \boldsymbol{e}_{p(2)} & \cdots & \boldsymbol{e}_{p(n)} \\ | & | & \cdots & | \end{bmatrix},$$

where p represents the corresponding permutation. Now note that the transpose can be written as

$$P = \begin{bmatrix} — \boldsymbol{e}_{p(1)}^t — \\ — \boldsymbol{e}_{p(2)}^t — \\ \vdots \\ — \boldsymbol{e}_{p(n)}^t — \end{bmatrix}.$$

Therefore, the ij^{th} element of the product P^tP is given by $\boldsymbol{e}_{p(i)}^t \boldsymbol{e}_{p(j)} = \delta_{p(i)p(j)} = \delta_{ij}$, meaning that $P^tP = \mathbb{I}$. We have used the fact that p is a bijection, so $p(i) = p(j)$ if and only if $i = j$. Thus, $P^{-1} = P^t$.

1.5 Cramer's Rule

Exercise 3. Let A be an $n \times n$ matrix with integer entries a_{ij}. Prove that A^{-1} has integer entries if and only if $\det A = \pm 1$.

Solution. First, suppose that $\det A = \pm 1$. If the entries of A^{-1} are b_{ij}, use

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A$$

to conclude that

$$b_{ij} = \frac{1}{\det A} (-1)^{i+j} \det A_{ji}.$$

Note that A_{ji} contains integer entries, hence its determinant must also be an integer via the complete expansion. Putting $\det A = \pm 1$ means that b_{ij} is always an integer.

Now suppose that A^{-1} has integer entries. Use $\det A = 1/\det A^{-1}$. Now both A and A^{-1} have integer entries, hence integer determinants, with $|\det A^{-1}| \geq 1$. This forces $\det A = \pm 1$.

Miscellaneous Problems

Exercise 2. Find a representation of the complex numbers by real 2 x 2 matrices which is compatible with addition and multiplication.

Solution. Consider the representation

$$z = a + ib \equiv \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Now, if $z = a + ib$, $w = c + id$, we have addition defined as

$$z + w \equiv \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} a+c & -b-d \\ b+d & a+c \end{bmatrix} \equiv (a+c) + i(b+d),$$

and multiplication as

$$zw \equiv \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac-bd & -ad-bc \\ ad+bc & ac-bd \end{bmatrix} \equiv (ac-bd) + i(ad+bc).$$

Finally,

$$|z|^2 = z\overline{z} = a^2 + b^2 = \det \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

Exercise 3. Find the Vandermonde determinant

$$\det A_n = \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{bmatrix}.$$

Solution. First look at the 2×2 case,

$$\det A_2 = \det \begin{bmatrix} 1 & 1 \\ a_1 & a_2 \end{bmatrix} = a_2 - a_1.$$

Now, look at the $n \times n$ case. Perform the row operations $R_k \to R_k - a_1 R_{k-1}$ for all rows $k = 2, \dots, n$. This leaves the determinant unchanged, so

$$\det A_2 = \det \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ 0 & a_2(a_2 - a_1) & a_3(a_3 - a_1) & \cdots & a_n(a_n - a_1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_2^{n-2}(a_2 - a_1) & a_3^{n-2}(a_3 - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{bmatrix}.$$

Using expansion by minors on the first column, we have

$$\det A_2 = \det \begin{bmatrix} a_2 - a_1 & a_3 - a_1 & \cdots & a_n - a_1 \\ a_2(a_2 - a_1) & a_3(a_3 - a_1) & \cdots & a_n(a_n - a_1) \\ \vdots & \vdots & \ddots & \vdots \\ a_2^{n-2}(a_2 - a_1) & a_3^{n-2}(a_3 - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{bmatrix}.$$

Factoring out $a_j - a_1$ from each j^{th} column gives

$$\det A_n = \prod_{j=2}^{n} (a_j - a_1) \times \det \begin{bmatrix} 1 & 1 & \cdots & 1 \\ a_2 & a_3 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_2^{n-2} & a_3^{n-2} & \cdots & a_n^{n-2} \end{bmatrix}$$

Continuing in this fashion, we get

$$\det A_n = \prod_{j=2}^{n} (a_j - a_1) \times \prod_{j=3}^{n} (a_j - a_2) \times \cdots \times (a_{n-1} - a_n).$$

This can be written down concisely as

$$\det A_n = \prod_{1 \le i < j \le n}^{n} (a_j - a_i)$$

Exercise 4. Consider a general system $AX = B$ of m linear equations in n unknowns. If the coefficient matrix A has a left inverse A', a matrix such that $A'A = \mathbb{I}_n$, then we may try to solve the system as follows.

$$\begin{aligned} AX &= B, \\ A'AX &= A'B \\ X &= A'B. \end{aligned}$$

But when we try to check our work by running the solution backward, we get into trouble:

$$\begin{aligned} X &= A'B \\ AX &= AA'B \\ AX &\stackrel{?}{=} B. \end{aligned}$$

We seem to want A' to be a right inverse: $AA' = \mathbb{I}_n$, which isn't what was given. Explain.

Solution. In the case that $m > n$, note that the left inverse is not necessarily unique. An example is

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbb{I}_2,$$

irrespective of a and b. Hence, $X = A'B$ is not unique, but rather is dependent on our choice of A'. If we had started with

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} p \\ q \\ r \end{bmatrix}$$

then we would have written

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} p + ar \\ q + br \end{bmatrix} = \begin{bmatrix} p \\ q \end{bmatrix} + r \begin{bmatrix} a \\ b \end{bmatrix}.$$

This means that the given argument is not sufficient to conclude $AA' = \mathbb{I}$.

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \end{bmatrix} = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 0 \end{bmatrix} \neq \mathbb{I}_3,$$

Note that this system is nonsense for $r \neq 0$ with no solutions, yet the left inverses A' do exist nonetheless. Here, $AX \neq B$ when $r \neq 0$, since

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} p + ar \\ q + ar \end{bmatrix} = \begin{bmatrix} p + ar \\ q + ar \\ 0 \end{bmatrix}.$$

In the case that $m < n$, A has no left inverse. Label the columns of A as A_i. Demanding $A'A = \mathbb{I}_n$ means that we want $A'A_i = \boldsymbol{e}_i$ for all $i = 1, \dots, n$. Since the $m \times n$ matrix A has more columns than rows, its columns must be linearly dependent, so without loss of generality, write the first column A_1 as a non-trivial linear combination of the rest,

$$A_1 = a_2 A_2 + a_3 A_3 + \cdots + a_n A_n.$$

Multiplying by A' gives

$$A'A_1 = \boldsymbol{e}_1 = a_2 \boldsymbol{e}_2 + a_3 \boldsymbol{e}_3 + \cdots + a_n \boldsymbol{e}_n,$$

which is a contradiction since the basis vectors $\{\boldsymbol{e}_i\}$ are linearly independent.

In the case $m = n$, it is indeed true that A' is also a right inverse of A. Note that if A'' is a right inverse of A with $AA'' = \mathbb{I}_n$, then

$$A' = A'\mathbb{I}_n = A'(AA'') = (A'A)A'' = \mathbb{I}_n A'' = A''.$$

To justify that A'' exists, note that $A'A = \mathbb{I}_n$ gives $\det A' \det A = 1$, so $\det A \neq 0$. Thus, A has full rank and its range must be the full n dimensional vector space of column vectors. Multiplying by A, we have $AA'A = A$ or $(AA' - \mathbb{I}_n)A = 0$. Recall that the range of A is the entire vector space, so $(AA' - \mathbb{I}_n)\boldsymbol{x} = \mathbf{0}$ for all possible column vectors $\boldsymbol{x}$. This forces $AA' - \mathbb{I}_n = 0$, or $AA' = \mathbb{I}_n$.

Exercise 5.

(a) Let A be a real 2×2 matrix, and let A_1, A_2 be the rows of A. Let P be the parallelogram whose vertices are 0, A_1, A_2, $A_1 + A_2$. Prove that the area of P is the absolute value of the determinant $\det A$ by comparing the effect of and elementary row operation on the area and on $\det A$.

(b) Prove an analogous result for $n \times n$ matrices.

Solution.

(a) First note that

$$\det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1,$$

which is consistent with the fact that the area of a unit square is 1. Now let a_{ij} be the elements of A. Perform the row operation which multiplies the top row by a_{11}, i.e. $R_1 \to a_{11}R_1$. We have

$$\det \begin{bmatrix} a_{11} & 0 \\ 0 & 1 \end{bmatrix} = a_{11}.$$

Now perform $R_1 \to R_1 + a_{12}R_2$. This gives

$$\det \begin{bmatrix} a_{11} & a_{12} \\ 0 & 1 \end{bmatrix} = a_{11}.$$

Next, perform $R_2 \to (a_{11}a_{22} - a_{12}a_{21})R_2$. This gives

$$\det \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{11}a_{22} - a_{12}a_{21} \end{bmatrix} = (a_{11}a_{22} - a_{12}a_{21})a_{11}.$$

Next, perform $R_2 \to R_2 + a_{21}R_1$. This gives

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{11}a_{21} & a_{11}a_{22} \end{bmatrix} = (a_{11}a_{22} - a_{12}a_{21})a_{11}.$$

Finally, perform $R_2 \to R_2/a_{11}$. This gives

$$\det A = \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

Note that if $a_{11} = 0$, we could have interchanged the roles of a_{11} and a_{21} at the beginning by interchanging the rows of A. This would have given the same result, up to a sign which we are not interested in. If both a_{11} and a_{22} are zero, note that the two rows are linearly dependent, with one being a multiple of the other, so the parallelogram they form has zero area. Thus, we have shown that any matrix A representing a parallelogram with non-zero area can be obtained from the identity matrix $\mathbb{I}_2$ by performing elementary row operations.

Now, we consider the effect of these row operations on the area of a parallelogram with legs A_1 and A_2. Note that the operation $A_1 \to kA_1$ for some real scaling factor k has the effect of scaling the area by the same factor k. The operation of interchanging the legs A_1 and A_2 has no effect on the area. The operation $A_1 \to A_1 + kA_2$ also has no effect on the area, because this has the effect of linearly shearing the parallelogram, in a manner parallel to the other leg A_2 which remains fixed. Thus, when we performed our row operations in the square to reach our parallelogram, our area transformed in precisely the same way as the unsigned determinant, which means that

$$\operatorname{area} A_{\|} = |\det A| = |a_{11}a_{22} - a_{12}a_{21}|.$$

(b) We use the fact that any square matrix A with non-zero determinant can be written as the product of row operations acting on the identity matrix $\mathbb{I}_n$, which represents the unit hypercube of hypervolume 1. The Gauss-Jordan elimination algorithm can be used to extract these operations. We see that all scaling operations will scale the hypervolume in the same way, all transpositions have no effect on the hypervolume, and all additions of linear combinations of other rows also have no effect, since they correspond to successive shearing of the hyperparallelopiped along a direction parallel to another leg. Thus, the area of the hypercube transformed in the same way as the unsigned determinant of A, so

$$\operatorname{hypervolume} A_{\|} = |\det A|.$$

Note that we are not interested in matrices with zero determinant, because such a matrix is not of full rank, hence its rows are linearly dependent. Thus, one of the legs of the corresponding hyperparallelopiped can be sheared until it is parallel to another, which immediately gives a zero hypervolume.

Exercise 6. Most invertible matrices can be written as a product $A = LU$ of a lower triangular matrix L and an upper triangular matrix U, where in addition all diagonal entries of U are 1.

(a) Prove uniqueness, that is, prove that there is at most one way to write A as a product.

(b) Explain how to compute L and U when the matrix A is given.

(c) Show that every invertible matrix can be written as a product LPU, where L, U are as above and P is a permutation matrix.

Solution. We first show that the determinant of a triangular matrix is equal to the product of its diagonals. To see this, note that this holds for all 2×2 lower triangular matrices,

$$\det \begin{bmatrix} a & 0 \\ c & d \end{bmatrix} = ad.$$

Next, suppose that this holds for all $n \times n$ lower triangular matrices. Using expansion of minors along the first row and our induction hypothesis on the minor A_{11}, compute

$$\det \begin{bmatrix} a_{11} & 0 & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & 0 & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} = a_{11} \det A_{11} + 0 + 0 + \cdots + 0 = a_{11}a_{22}a_{33} \cdots a_{nn}.$$

For upper triangular matrices, simply note that $\det U = \det U^t$, and U^t is lower triangular.

Now, we show that the inverse of a triangular matrix is also triangular of the same kind. Note that this holds for all invertible 2×2 matrices, with

$$\begin{bmatrix} a & 0 \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad} \begin{bmatrix} d & 0 \\ -c & a \end{bmatrix}.$$

Next, suppose that an invertible lower triangular matrix L has an inverse L^{-1}, whose columns are labelled $\boldsymbol{x}_j$. Since $LL^{-1} = \mathbb{I}_n$, we want

$$L\boldsymbol{x}_j = \boldsymbol{e}_j.$$

We claim that $(x_j)_i = 0$ for all $i < j$. To see this, note that the first $j-1$ rows expand to

$$\begin{aligned} 0 &= L_{11}x_{j1} \\ 0 &= L_{12}x_{j1} + L_{22}x_{j2} \\ \vdots & \quad \vdots \\ 0 &= L_{j-1,1}x_{j,j-1} + \cdots + L_{j-1,j-1}x_{j,j-1} \end{aligned}$$

All L_{ij} with $i < j$ are zero, and L_{ii} are non-zero since L is invertible hence $\det L \neq 0$. Thus, the first equation gives $x_{j1} = 0$, which when plugged into the second gives $x_{j2} = 0$, and so on up to $x_{j,j-1} = 0$. Hence, $L^{-1}_{ij} = 0$ for all $i < j$, making it a lower triangular matrix. In addition, the j^{th} row reads

$$1 = L_{j1}x_{j1} + \cdots + L_{j,j-1}x_{j,j-1} + L_{jj}x_{jj}.$$

All terms but the last one are 0, so the diagonal elements satisfy $L_{jj}L^{-1}_{jj} = 1$. Like before, for a lower triangular matrix U, use $(U^t)^{-1} = (U^{-1})^t$.

Finally, the product of two triangular matrices of the same kind give another triangular matrix of the same kind. Suppose that A and B are two lower triangular matrices. The ij^{th} element of their product AB is

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Now, $a_{ik} = 0$ for all $i < k$ and $b_{ki} = 0$ for all $k < j$. Thus, when $i < j$, we have $c_{ij} = 0$, hence AB is also lower triangular. Again for upper triangular matrices X, Y, use $(XY)^t = Y^tX^t$.

(a) Suppose that $A = LU = L'U'$ are two LU decompositions of A. Note that $\det A \neq 0$ from its invertibility, hence L, U, L', U' are all invertible. This gives

$$L^{-1}LU = L^{-1}L'U', \qquad U = L^{-1}L'U', \qquad U(U')^{-1} = L^{-1}L'.$$

Now, the left side is upper triangular while the right side is left triangular. Also, the left side has all 1's along its diagonal. This forces

$$U(U')^{-1} = \mathbb{I}_n = L^{-1}L', \qquad U = U', \quad L = L'.$$

(b) The elements of L and U can be obtained by brute force, solving the system $A = LU$ with $n(n+1)/2 + (n-1)n/2 = n^2$ unknowns.

(c) Note that after performing Gaussian elimination on an invertible matrix A, we are left with an upper triangular matrix U with 1's along its diagonal. Also, each elementary operation we performed can be represented by a lower triangular matrix. This is because all scaling matrices are diagonal, and in all cases where we added one row to another we always added higher row to ones lower down. Thus, the product of all these elementary matrices is a lower triangular matrix L, which means $LA = U$. This gives the desired decomposition, $A = L^{-1}U$.

However, we may have to exchange rows while performing the elimination process, which happens when one of the diagonal elements becomes zero. By performing this permutation of rows at the very end, we have actually decomposed $PLA = U$. The inverse of a permutation is another permutation, hence we have the desired decomposition $A = L^{-1}P^{-1}U$.

Exercise 7. Consider a system of n linear equations in n unknowns: $AX = B$, where A and B have *integer* entries. Prove or disprove the following.

(a) The system has a rational solution if $\det A \neq 0$.

(b) If the system has a rational solution, then it also has an integer solution.

Solution.

(a) If $\det A \neq 0$, then A is invertible. Since A has integer entries, its determinant is an integer and its adjoint has integer entries, which means that $A^{-1} = (\operatorname{adj} A)/\det A$ has rational entries. Also, B has integer entries so the solution $X = A^{-1}B$ must also be rational.

(b) This is false. Consider the system

$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

This has the unique solution $x = y = \frac{1}{2}$.

Exercise 8. Let A, B be $m \times n$ and $n \times m$ matrices. Prove that $\mathbb{I}_m - AB$ is invertible if and only if $\mathbb{I}_n - BA$ is invertible.

Solution. Note that

$$B(\mathbb{I}_m - AB) = B - BAB = (\mathbb{I}_n - BA)B,$$
$$A(\mathbb{I}_n - BA) = A - ABA = (\mathbb{I}_m - AB)A.$$

Set $X = \mathbb{I}_m - AB$, $Y = \mathbb{I}_n - BA$, whence

$$BX = YB, \qquad AY = XA.$$

First suppose that X is invertible. If A is invertible, then $AY = XA$ gives $Y = A^{-1}XA$, so we can check that $Y^{-1} = A^{-1}X^{-1}A$.

$$(A^{-1}X^{-1}A)Y = A^{-1}X^{-1}A\,A^{-1}XA = \mathbb{I}_n.$$

If A is not invertible but B is invertible, then use $BX = YB$ to write $Y = BXB^{-1}$, so we can check that $Y^{-1} = BX^{-1}B^{-1}$.

$$(BX^{-1}B^{-1})Y = BX^{-1}B^{-1}\,BXB^{-1} = \mathbb{I}_n.$$

Now suppose that neither A nor B is invertible. Consider the products

$$(\mathbb{I}_n + BX^{-1}A)Y = Y + BX^{-1}AY = Y + BX^{-1}XA = Y + BA = \mathbb{I}_n,$$

$$Y(\mathbb{I}_n + BX^{-1}A) = Y + YBX^{-1}A = Y + BXX^{-1}A = Y + BA = \mathbb{I}_n.$$

Thus, $Y^{-1} = \mathbb{I}_n + BX^{-1}A$.