5. MATRICES AND DETERMINANTS

§5.1. Matrices

The previous three chapters gave a strong geometric motivation for studying matrices and determinants up to dimension 3. But matrices are an important tool in many non-geometric areas. While a knowledge of the previous chapters may be useful they are not necessary. You can begin with this chapter.

Often data is best arranged in a table of rows and columns. Such a table can be considered as a single mathematical object, called a matrix (plural 'matrices'). An $m \times n$ **matrix** is a rectangular array of numbers arranged in m rows and n columns. The numbers can come from any field, such as the field of real numbers or the field of complex numbers. The entries in the table are called the **components** of the matrix.

We use a double suffix notation to represent the components, with the first suffix representing the row and the second representing the column. So, for example, we might denote the *i-j* component (*i*'th row, *j*'th column) by $a_{i,j}$. Usually we leave out the comma and write the entry in the second row and third column as a_{23}. With larger matrices there's possible confusion with this system – is a_{123} the entry in the 12th row and 3rd column or the one in the 1st row and 23rd column? In such cases we'd need to reinstate the comma.

So a typical $m \times n$ matrix would look like this:

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots\ldots & a_{1n} \\ a_{21} & a_{22} & \ldots\ldots & a_{2n} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_{m1} & a_{m2} & \ldots\ldots & a_{mn} \end{pmatrix}$$

If $m = n$ we call A a **square matrix**. A useful shorthand is to write **A = (a_{ij}).** So the 12 × 12 matrix (a_{ij}), where a_{ij} is i times j, is simply the standard multiplication table.

1	2	3	4	5	6	7	8	9	10	11	12
2	4	6	8	10	12	14	16	18	20	22	24
3	6	9	12	15	18	21	24	27	30	33	36
4	8	12	16	20	24	28	32	36	40	44	48
5	10	15	20	25	30	35	40	45	50	55	60
6	12	18	24	30	36	42	48	54	60	66	72
7	14	21	28	35	42	49	56	63	70	77	84
8	16	24	32	40	48	56	64	72	80	88	96
9	18	27	36	45	54	63	72	81	90	99	108
10	20	30	40	50	60	70	80	90	100	110	120
11	22	33	44	55	66	77	88	99	110	121	132
12	24	36	48	60	72	84	96	108	120	132	144

If A is an $m \times n$ matrix and we swap rows and columns we get an $n \times m$ matrix called the **transpose** of A. We write this as $\mathbf{A^T}$ and we can write $A^T = (a_{ji})$, meaning that the *i-j* component is now a_{ji}, the component that was in the *j-i* position.

$$\text{In other words } A^T = \begin{pmatrix} a_{11} & a_{21} & \ldots\ldots & a_{m1} \\ a_{12} & a_{22} & \ldots\ldots & a_{m2} \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ a_{1n} & a_{2n} & \ldots\ldots & a_{mn} \end{pmatrix}$$

Clearly $A^{TT} = A$. If $A^T = A$, as in the multiplication table, we say that A is a **symmetric matrix**. It is symmetric about the diagonal that goes from top left to bottom right corners. (The other diagonal has no significance for matrices, so when we say *the* diagonal you know which one we mean.)

Example 1: If $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ then $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

The **trace** of a square matrix is the sum of the diagonal components. If $A = (a_{ij})$ is an $n \times n$ matrix then $\mathbf{tr}\boldsymbol{A} = a_{11} + a_{22} + \ldots + a_{nn}$. (I suppose there's no reason why trace can't be defined for non-square matrices. It's just that it doesn't seem to be a useful concept unless the matrix is square.)

Example 2: If A = $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ then trA = 1 + 5 + 9 = 15.

A **diagonal matrix** is a square matrix where every component that lies off the diagonal is zero. It has the form:

$$\begin{pmatrix} a_{11} & 0 & \ldots\ldots & 0 \\ 0 & a_{22} & \ldots\ldots & 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ 0 & 0 & \ldots\ldots & a_{nn} \end{pmatrix}$$

A **scalar matrix** is a diagonal matrix where all the diagonal components are equal.
It has the form:

$$\begin{pmatrix} x & 0 & \ldots\ldots & 0 \\ 0 & x & \ldots\ldots & 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ 0 & 0 & \ldots\ldots & x \end{pmatrix}$$

The special cases of scalar matrices are:

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \ldots\ldots & 0 \\ 0 & 0 & \ldots\ldots & 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ 0 & 0 & \ldots\ldots & 0 \end{pmatrix} \text{ and}$$

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \ldots\ldots & 0 \\ 0 & 1 & \ldots\ldots & 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ 0 & 0 & \ldots\ldots & 1 \end{pmatrix}$$

A **column vector** is an $m \times 1$ matrix written as $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_m \end{pmatrix}$.

It is customary to print vectors in bold type to distinguish them from scalars (ordinary field elements). A **row vector** is a $1 \times n$ matrix $\mathbf{v}^T = (x_1, x_2, \ldots, x_n)$.

Note that matrices in general can be considered to be row or column vectors if we allow their components to be the columns, or rows, respectively. So we could write an $m \times n$ matrix as $A = (\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_n)$. This would mean that the columns of A are the $\mathbf{c}_i$.

Then A^T would be the column vector $\begin{pmatrix} \mathbf{c}_1^T \\ \mathbf{c}_2^T \\ \ldots \\ \mathbf{c}_n^T \end{pmatrix}$.

Or we could write $A = \begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \ldots \\ \mathbf{r}_m \end{pmatrix}$ where the $\mathbf{r}_i$ are the rows.

In that case $A^T = (\mathbf{r}_1^T, \mathbf{r}_2^T, \ldots, \mathbf{r}_m^T)$. Never mind that we said that the components of a matrix are supposed to come from a field. We can break this rule sometimes, but don't expect all the theory of matrices over a field to apply if the components are vectors!

Example 3: If A = $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ then A = (**u**, **v**, **w**) where

$$\mathbf{u} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} 3 \\ 6 \end{pmatrix}.$$

Also A = $\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}$ where **a** = (1, 2, 3) and **b** = (4, 5, 6).

Addition of matrices is defined in the obvious way. If A = (a_{ij}) and B = (b_{ij}) are two $m \times n$ matrices then:

A + B is the $m \times n$ matrix ($a_{ij} + b_{ij}$).

That is, we simply add corresponding components. Addition is clearly associative and commutative. The zero matrix **0** is the identity and –A, meaning ($-a_{ij}$), is the additive inverse of A.

Example 4: If A = $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ and B = $\begin{pmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \end{pmatrix}$ then A + B = $\begin{pmatrix} 3 & 5 & 7 \\ 9 & 11 & 13 \end{pmatrix}$, –A = $\begin{pmatrix} -1 & -2 & -3 \\ -4 & -5 & -6 \end{pmatrix}$ and B – A, meaning B + (– A), is $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

Note that it's not possible to add two matrices with different dimensions. So in example 1 we can't add A to A^T.

Multiplication by a scalar is equally obvious. If λ is an element of whatever field we're working with and A = (a_{ij}) then λA = (λa_{ij}). We just multiply each component by the scalar. Clearly every scalar matrix has the form λI for some λ.

Example 5: A = $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ then 2A = $\begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}$.

Addition and scalar multiplication interact with transpose and trace in a natural way.

$(A + B)^T = A^T + B^T$;
$(\lambda A)^T = \lambda A^T$;
$\mathrm{tr}(A + B) = \mathrm{tr}A + \mathrm{tr}B$ and
$\mathrm{tr}(\lambda A) = \lambda\mathrm{tr}A$.

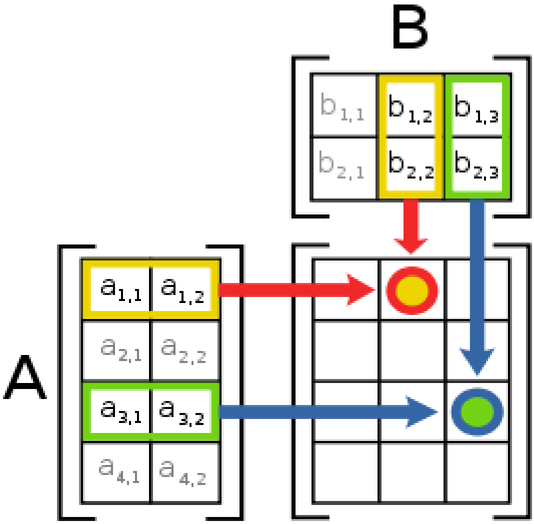
§5.2. Matrix Multiplication

By contrast, **matrix multiplication** is far less obvious. If A = (a_{ij}) is an $m \times n$ matrix and B = (b_{ij}) is an $n \times r$ matrix then AB = C = (c_{ij}) where

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} \text{ for each } i, j.$$

The easiest way to remember this is as follows: to get the *i*-*j* component of the product we go along the *i*'th row of A and down the *j*'th column of B, multiplying corresponding components and adding these products.

That's why it's important for the number of columns in the first factor to be the same as the number of rows in the second – so that we don't run out of one lot of factors before we've finished with the other.



In the following theorems we assume that the matrices have compatible sizes for the products to be defined.

Theorem 1 (Associative Law): (AB)C = A(BC).
Proof: Let AB = M and let BC = N.
The *i*-*j* component of (AB)C = MC is

$$\sum_k m_{ik}c_{kj} = \sum_k \left(\sum_t a_{it}c_{tk}\right)c_{kj} = \sum_k \sum_t a_{it}b_{tk}c_{kj}.$$

The *i*-*j* component of A(BC) = AN is

$$\sum_t a_{it} n_{tj} = \sum_t a_{it} \left(\sum_k b_{tk} c_{kj} \right) c_{kj} = \sum_t \sum_k a_{it} b_{tk} c_{kj}.$$

These are equal since the order of summation may be interchanged.

Example 6: If A = $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$, B = $\begin{pmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}$ and C = $\begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$

$\begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix}$ find, where possible:

(i) A + B; (ii) B + B^T; (iii) 2A;
(iv) AB; (v) BA; (vi) AC;
(vii) CA; (viii) A^2; (ix) B^2; (x) AA^TC.

Solution: (i) undefined;

(ii) B + B^T = $\begin{pmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} + \begin{pmatrix} 9 & 6 & 3 \\ 8 & 5 & 2 \\ 7 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 18 & 14 & 10 \\ 14 & 10 & 6 \\ 10 & 65 & 2 \end{pmatrix}$;

(iii) 2A = $\begin{pmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \end{pmatrix}$;

(iv) AB = $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 30 & 24 & 18 \\ 84 & 69 & 54 \end{pmatrix}$;

(v) undefined;

(vi) undefined;

(vii) CA = $\begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 18 & 24 & 30 \\ 38 & 52 & 48 \end{pmatrix}$;

(viii) undefined;

(ix) $B^2 = \begin{pmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 150 & 126 & 102 \\ 96 & 81 & 66 \\ 42 & 36 & 30 \end{pmatrix}$;

(x) $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 14 & 22 \\ 32 & 77 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} = \begin{pmatrix} 160 & 232 \\ 526 & 744 \end{pmatrix}$.

Theorem 2 (Distributive Laws):

A(B + C) = AB + AC and (A + B)C = AC + BC.

Proof: The *i*-j component of A(B + C) is

$$\sum_k a_{ik}(b_{kj} + c_{kj}) = \sum_k a_{ik}b_{kj} + \sum_k a_{ik}c_{kj}$$

which is the *i*-*j* component of AB + AC.

Similarly we can prove that (A + B)C = AC + BC for all matrices A, B, C of compatible sizes.

Theorem 3: $(AB)^T = B^TA^T$.

Proof: The *i*-*j* component of $(AB)^T$ is the *j*-*i* component of AB = $\sum_k a_{jk}b_{ki} = \sum_k b_{ki}a_{jk}$

$= \sum_k$(*i*-*k* component of B^T)(*k*-*j* component of A^T)

= *i*-*j* component of B^TA^T.

Theorem 4: tr(AB) = tr(BA).

Proof: $\text{tr}(AB) = \sum_i \left(\sum_k a_{ik} b_{ki} \right)$

$$= \sum_k \left(\sum_i a_{ik} b_{ki} \right)$$

$$= \sum_k \left(\sum_i b_{ki} a_{ik} \right)$$

$$= \sum_k (k\text{-}k \text{ component of BA}) \,)$$

$$= \text{tr}(BA).$$

NOTE: It is *not* the case that tr(AB) = trA.trB.

As we saw with 2 × 2 matrices matrix multiplication does not behave as nicely as multiplication in a field. The commutative law fails and it is possible for the product of two non-zero matrices to be zero.

A square matrix A is **invertible** (or 'non-singular') if it has an inverse under multiplication. The inverse B has to be 2-sided, meaning that not only is AB = I but BA = I. Some matrices have left inverses but no right inverse, or vice versa. But if a matrix has both a left and a right inverse they are equal.

Theorem 5: If A is invertible the inverse is unique.
Proof: Suppose B_1 and B_2 are inverses for A.
Then $B_1 = B_1(AB_2) = (B_1A)B_2 = B_2$.

You can see how it is important for the inverse to work on both sides. Matrices with one-sided inverses can have more than one.

Example 7: Since $\begin{pmatrix} 1 & 0 & x \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ for all x, $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ has infinitely many left inverses. It must therefore have no right inverses.

Theorem 6: If A, B are invertible then so is AB and

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Proof: $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1}$.
Similarly $(B^{-1}A^{-1})(AB) = I$.

§5.3. Determinants

The **determinant** of a square matrix is a number that, among other things, determines whether or not the matrix is invertible. We define it inductively by defining $n \times n$ determinants in terms of certain $(n-1) \times (n-1)$ determinants obtained by deleting a row and a column. We denote the determinant of A by **|A|**, but if we're listing the components of the matrix we omit the matrix parentheses. So we write $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$ instead of $\left|\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right|$.

We define $\boldsymbol{\delta^{i}_{j}}$**(A)** to be the matrix obtained from A by deleting the i'th row and the j'th column. Of course this won't work if we're trying to define the determinant of 1×1 matrices, but there we define the determinant to be the only component that the matrix has. So we define $|(a)| = a$. Do not confuse this with absolute value.

If $n \geq 2$ and A is an $n \times n$ matrix we define

$$|\mathrm{A}| = \sum_{j} (-1)^{1+j}\, \mathrm{a}_{1j}\, |\delta^{1}_{j}(\mathrm{A})| \, .$$

We call this expansion the **First Order expansion**. It's obvious from the definition that if the 1st row is zero the determinant is zero.

Of course we want this definition to agree with the one we gave in earlier chapters for 2×2 and 3×3 matrices. Let's check that.

For 2×2 determinants we have

$$\begin{vmatrix} a & b \\ c & \mathrm{d} \end{vmatrix} = a|(d)| - b|(c)| = ad - bc.$$

For 3×3 determinants

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \, .$$

The determinant of a diagonal matrix is clearly the product of the diagonal components. This is because only the first term in the first order expansion is non-zero. For other determinants the amount of work can be considerable. Just imagine working out a 10 × 10 determinant. This involves ten 9 × 9 determinants, each requiring nine 8 × 8 determinants, and so on. An analysis of the process shows that there will be over 6 million multiplications!

Example 8:

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 3 & 4 & 1 \end{vmatrix} - 2\begin{vmatrix} 4 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 4 & 1 \end{vmatrix} + 3\begin{vmatrix} 4 & 1 & 3 \\ 3 & 4 & 2 \\ 2 & 3 & 1 \end{vmatrix} - 4\begin{vmatrix} 4 & 1 & 2 \\ 3 & 4 & 1 \\ 2 & 3 & 4 \end{vmatrix}.$$

Now $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 1 & 2 \\ 3 & 4 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 4 & 1 \end{vmatrix} - 2\begin{vmatrix} 4 & 2 \\ 3 & 1 \end{vmatrix} + 3\begin{vmatrix} 4 & 1 \\ 3 & 4 \end{vmatrix}$

$= (1 - 8) - 2(4 - 6) + 3(16 - 3) = -7 + 4 + 39 = 36.$

$$\begin{vmatrix} 4 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 4 & 1 \end{vmatrix} = 4\begin{vmatrix} 1 & 2 \\ 4 & 1 \end{vmatrix} - 2\begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} + 3\begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix}$$

$= 4(1 - 8) - 2(3 - 4) + 3(12 - 2) = -28 + 2 + 30 = 4.$

$$\begin{vmatrix} 4 & 1 & 3 \\ 3 & 4 & 2 \\ 2 & 3 & 1 \end{vmatrix} = 4\begin{vmatrix} 4 & 2 \\ 3 & 1 \end{vmatrix} - \begin{vmatrix} 3 & 2 \\ 2 & 1 \end{vmatrix} + 3\begin{vmatrix} 3 & 4 \\ 2 & 3 \end{vmatrix}$$

$= 4(4 – 6) – (3 – 4) + 3(9 – 8) = – 8 + 1 + 3 = – 4.$

$$\begin{vmatrix} 4 & 1 & 2 \\ 3 & 4 & 1 \\ 2 & 3 & 4 \end{vmatrix} = 4\begin{vmatrix} 4 & 1 \\ 3 & 4 \end{vmatrix} - \begin{vmatrix} 3 & 1 \\ 2 & 4 \end{vmatrix} + 2\begin{vmatrix} 3 & 4 \\ 2 & 3 \end{vmatrix}$$

$= 4(16 – 3) – (12 – 2) + 2(9 – 8) = 52 – 10 + 2 = 44.$

Hence the final answer is $36 – 2(4) + 3(–4) – 4(44)$
$= 36 – 8 – 12 – 176 = – 160.$

It’s clear that working out determinants can be very labour intensive, and prone to error. However there are several tricks and shortcuts. This is particularly so when the matrix has a certain pattern. Let me let you into a secret. With the above example you’ll notice that it follows a very simple pattern – each row changes to the next by moving to the right, with the end components coming back to the left. Because of this I could work out this determinant quite quickly. (You’ll be able to do this later.) Then I worked it out the long way, and kept making silly arithmetic mistakes for a little while. Luckily I knew in advance what it should be!

We now develop a number of very nice properties of determinants, properties that can make it a lot easier to evaluate a determinant. To assist us in proving these we introduce a notation for the matrix you get by deleting two rows and two columns. This leads to another formula for a determinant. This can often reduce the amount of computation in working out a determinant, but it's main use is in proving the properties.

If A is a square matrix then $\boldsymbol{\delta_{ij}^{st}(A)}$ is the matrix obtained from A by deleting rows *s*, *t* and columns *i*, *j*.

Theorem 7 (Second Order Expansion):

$$|A| = \sum_{j} \sum_{i<j} (-1)^{1+i+j} \begin{vmatrix} a_{1i} & a_{1j} \\ a_{2i} & a_{2j} \end{vmatrix} .|\delta_{ij}^{12}(A)|) .$$

Proof: $|A| = \sum_{i<j} (-1)^{1+j} |\delta_{j}^{1}(A)|) =$

$$\sum_{j} (-1)^{1+\mathrm{i}} \sum_{i<j} \left\{ (-1)^{1+i} a_{2i} |\delta_{ij}^{12}(A)| \right.$$

$$\left. + \sum_{i>j} (-1)^{i} a_{2i} |\delta_{ij}^{12}(A)| \right\}$$

(when $i > j$ the j'th row being deleted moves the i'th row up one place)

$$= \sum_{i<j} (-1)^{1+i+j} a_{1j}a_{2i} |\delta_{ij}^{12}(A)|$$

$$+ \sum_{i>j} (-1)^{1+i+j} a_{1i}a_{2j} |\delta_{ij}^{12}(A)|$$

(the dummy variables i, j were interchanged in the 2nd sum)

$$= \sum_{i<j} (-1)^{1+i+j} \begin{vmatrix} a_{1i} & a_{1j} \\ a_{2i} & a_{2j} \end{vmatrix} . |\delta_{ij}^{12}(A)|) .$$

Example 9: Evaluate the determinant $\begin{vmatrix} 6 & 2 & 1 & -3 \\ 4 & 0 & 5 & 3 \\ 2 & -4 & 6 & 1 \\ 5 & 1 & 0 & -4 \end{vmatrix}$

(i) by the first order expansion;
(ii) by the second order expansion.

Solution:

(i) $$\begin{vmatrix} 6 & 2 & 1 & -3 \\ 4 & 0 & 5 & 3 \\ 2 & -4 & 6 & 1 \\ 5 & 1 & 0 & -4 \end{vmatrix} = 6\begin{vmatrix} 0 & 5 & 3 \\ -4 & 6 & 1 \\ 1 & 0 & -4 \end{vmatrix} - 2\begin{vmatrix} 4 & 5 & 3 \\ 2 & 6 & 1 \\ 5 & 0 & -4 \end{vmatrix} + \begin{vmatrix} 4 & 0 & 3 \\ 2 & -4 & 1 \\ 5 & 1 & -4 \end{vmatrix} + 3\begin{vmatrix} 4 & 0 & 5 \\ 2 & -4 & 6 \\ 5 & 1 & 0 \end{vmatrix}$$

$$= -30\begin{vmatrix} -4 & 1 \\ 1 & -4 \end{vmatrix} + 18\begin{vmatrix} -4 & 6 \\ 1 & 0 \end{vmatrix} - 8\begin{vmatrix} 6 & 1 \\ 0 & -4 \end{vmatrix} + 10\begin{vmatrix} 2 & 1 \\ 5 & -4 \end{vmatrix} - 6\begin{vmatrix} 2 & 6 \\ 5 & 0 \end{vmatrix}$$
$$+ 4\begin{vmatrix} -4 & 1 \\ 1 & -4 \end{vmatrix} + 3\begin{vmatrix} 2 & -4 \\ 5 & 1 \end{vmatrix} + 12\begin{vmatrix} -4 & 6 \\ 1 & 0 \end{vmatrix} + 15\begin{vmatrix} 2 & -4 \\ 5 & 1 \end{vmatrix}$$

$= (-30)(15) + (18)(-6) - (8)(-24) + (10)(-13) - (6)(-30)$
$+ (4)(15) + (3)(22) + (12)(-6) + (15)(22)$
$= 68.$

(ii) $\begin{vmatrix} 6 & 2 & 1 & -3 \\ 4 & 0 & 5 & 3 \\ 2 & -4 & 6 & 1 \\ 5 & 1 & 0 & -4 \end{vmatrix}$

$= \begin{vmatrix} 6 & 2 \\ 4 & 0 \end{vmatrix} \cdot \begin{vmatrix} 6 & 1 \\ 0 & -4 \end{vmatrix} - \begin{vmatrix} 6 & 1 \\ 4 & 5 \end{vmatrix} \cdot \begin{vmatrix} -4 & 1 \\ 1 & -4 \end{vmatrix} + \begin{vmatrix} 6 & -3 \\ 4 & 3 \end{vmatrix} \cdot \begin{vmatrix} -4 & 6 \\ 1 & 0 \end{vmatrix}$

$+ \begin{vmatrix} 2 & 1 \\ 0 & 5 \end{vmatrix} \cdot \begin{vmatrix} 2 & 1 \\ 5 & -4 \end{vmatrix} - \begin{vmatrix} 2 & -3 \\ 0 & 3 \end{vmatrix} \cdot \begin{vmatrix} 2 & 6 \\ 5 & 0 \end{vmatrix} + \begin{vmatrix} 1 & -3 \\ 5 & 3 \end{vmatrix} \cdot \begin{vmatrix} 2 & -4 \\ 5 & 1 \end{vmatrix}$

$= (-8)(-24) - (26)(15) + (30)(-6) + (10)(-13)$
$- (6)(-30) + (18)(22)$

$= 68.$

For a 3 × 3 determinant the Second Order Expansion actually requires more work so you would only use it for 4 × 4 determinants and above. For a 10 × 10 determinant the Second Order Expansion requires just over 600,000 multiplications compared with the First Order Expansion which takes over 6 million. However in practice one uses neither. There are far more efficient techniques. But to develop these techniques we need to know certain properties of determinants, and these require both the First and Second Order Expansions to prove.

Theorem 8: Interchanging rows i and j of a determinant changes its sign.

Proof: Suppose we have an $n \times n$ determinant. We prove the theorem by induction on n.

It's clearly true if $n = 2$. Suppose $n \geq 3$ and suppose $j > i$.

Case I: *i* ≥ 2: This follows by induction from the 1st Order Expansion (the definition of determinants).
Case II: *i* = 1, *j* = 2: This follows from the 2 × 2 case and the 2nd Order Expansion.
Case III: *i* = 1, *j* > 2: We can swap the 1st and *j*'th rows by swapping the 1st and 2nd, then the 2nd and *j*'th and finally swapping the 1st and 2nd again.

$$\begin{matrix} R_1 & & R_2 & & R_2 & & R_j \\ R_2 & \to & R_1 & \to & R_j & \to & R_2 \\ R_j & & R_j & & R_1 & & R_1 \end{matrix}.$$

Each swap changes the sign and three changes of sign are equivalent to one change of sign.
Corollary: If a matrix has a row of zeros its determinant is zero.
Proof: Just swap that row with the first.

Example 10: Since $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{vmatrix} = -160$, $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 2 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix} = 160.$

$(R_2 \leftrightarrow R_4)$

Theorem 9: Multiplying (dividing) row *i* by λ multiplies (divides) the determinant by λ.
Proof: Suppose we have an $n \times n$ determinant. We prove the theorem by induction on *n*.
It's clearly true if $n = 1$. Suppose $n \geq 2$.
Case I: *i* = 1: This follows directly from the 1st order expansion.

Case II: *i* > 2: This follows by induction from the 1st order expansion.

Theorem 10: Adding (subtracting) λ times the *j*'th row to (from) the *i*'th does not change the determinant.

Proof: Suppose we have an $n \times n$ determinant. We prove the theorem by induction on n.

It's clearly true for $n = 2$. Suppose $n \geq 3$.

Case I: *j* > *i* ≥ 2: This follows by induction from the 1st Order expansion (the definition of determinants).

Case II: *i* = 1, *j* = 2: This follows from the 2 × 2 case and the 2nd Order Expansion.

Case III: *i* = 1, *j* > 2: We can add λ times the *j*'th row to the 1st by swapping the 2nd and *j*'th, then adding (subtracting) λ times the new 2nd row to (from) the 1st and then swapping the 2nd and *j*'th rows again.

$$\begin{matrix} R_1 \\ R_2 \\ R_j \end{matrix} \rightarrow \begin{matrix} R_1 \\ R_j \\ R_2 \end{matrix} \rightarrow \begin{matrix} R_1+\lambda R_j \\ R_j \\ R_2 \end{matrix} \rightarrow \begin{matrix} R_1+\lambda R_j \\ R_2 \\ R_j \end{matrix}.$$

Each of the two swaps changes the sign while the middle operation makes no change.

Corollary: If a matrix has two identical rows (or two identical columns) its determinant is zero.

Proof: Subtracting one of these two rows from the other doesn't change the determinant, yet results in a matrix with a zero row.

Example 11:

Since $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{vmatrix} = -160$, $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & -2 & -2 & 2 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{vmatrix} = -160$. $(R_2 - R_4)$

These properties can be used to simplify determinants, and so reduce the effort in evaluating them.

Example 12: Evaluate $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 1002 & 2003 & 3004 & 4002 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix}$.

Solution: Subtracting 1000 times row 1 from row 2,

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 1002 & 2003 & 3004 & 4002 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 2 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{vmatrix} = 160$$

(from Example 10).

Example 13: Evaluate $\begin{vmatrix} 6 & 2 & 1 & 5 \\ 2 & 7 & 2 & 1 \\ 3 & 5 & 0 & 2 \\ 6 & 1 & 3 & 9 \end{vmatrix}$.

Solution: $\begin{vmatrix} 6 & 2 & 1 & 5 \\ 2 & 7 & 2 & 1 \\ 3 & 5 & 0 & 2 \\ 6 & 1 & 3 & 9 \end{vmatrix} = \begin{vmatrix} 6 & 2 & 1 & 5 \\ -10 & 3 & 0 & -9 \\ 3 & 5 & 0 & 2 \\ -12 & -5 & 0 & -6 \end{vmatrix}$.

$(R_2 - 2R_1$ and $R_4 - 3R_1)$

Doing a 1st Order Expansion three of the four 3 × 3 determinants are zero because they will have a column of zeros.

So $\begin{vmatrix} 6 & 2 & 1 & 5 \\ 2 & 7 & 2 & 1 \\ 3 & 5 & 0 & 2 \\ 6 & 1 & 3 & 9 \end{vmatrix} = (-1)^2 \begin{vmatrix} -10 & 3 & -9 \\ 3 & 5 & 2 \\ -12 & -5 & -6 \end{vmatrix}$

$$= \begin{vmatrix} -10 & 3 & -9 \\ 3 & 5 & 2 \\ -12 & -5 & -6 \end{vmatrix}$$

$$= \begin{vmatrix} -10 & 3 & -9 \\ 23 & -1 & 22 \\ -12 & -5 & -6 \end{vmatrix} \quad (R_2 - 2R_2)$$

$$= \begin{vmatrix} 23 & -1 & 22 \\ -10 & 3 & -9 \\ -12 & -5 & -6 \end{vmatrix} \quad (R_1 \leftrightarrow R_2)$$

$$= \begin{vmatrix} 23 & -1 & 22 \\ 59 & 0 & 47 \\ -127 & 0 & -116 \end{vmatrix}$$

($R_2 + 3R_1$ and $R_3 - 5R_1$)

Expanding we note that 2 of the 3 2×2 determinants will have a column of zeros, and so be a zero determinant.

Hence $\begin{vmatrix} 6 & 2 & 1 & 5 \\ 2 & 7 & 2 & 1 \\ 3 & 5 & 0 & 2 \\ 6 & 1 & 3 & 9 \end{vmatrix} = \begin{vmatrix} 59 & 47 \\ -127 & -116 \end{vmatrix}$

$$= -(59)(116) + (127)(47) = -875.$$

§5.4. The Vandermonde Determinant

An important special case of a determinant is one of the form

$$\begin{vmatrix} x_1 & x_1^2 & \dots & x_1^n \\ x_2 & x_2^2 & \dots & x_2^n \\ \dots & \dots & \dots & \dots \\ x_n & x_n^2 & \dots & x_n^n \end{vmatrix}.$$

It is called a **Vandermonde determinant** and is denoted by **V(x_1, x_2, …, x_n)**.

The Vandermonde determinant is actually a polynomial in the n commuting variables $x_1, x_2, \dots, x_n$. The total degree (that is, the sum of the degrees in each of the variables) is clearly $1 + 2 + \dots + n = ½\, n(n + 1)$.

Theorem 11:

$$V(x_1, x_2, \dots, x_n) = \prod_{i<j} (x_j - x_i)\, x_1 x_2 \dots x_n,$$

Where the product is over all pairs (i, j) with $i < j$.

Proof: If $x_i = x_j$, for $i < j$, then $V(x_1, x_2, \dots, x_n) = 0$ because it will have two equal rows.

Hence $x_i - x_j$ is a factor of $V(x_1, x_2, \dots, x_n)$ for all $i < j$.

Also, if any of the x_i is zero then $V(x_1, x_2, \dots, x_n) = 0$ and so $x_1, x_2, \dots, x_n$ are factors.

There are $½\, n(n - 1) + n = ½\, n(n + 1)$ of these factors and since they're all coprime as polynomials their product must divide $V(x_1, x_2, \ldots, x_n)$. But the total degree of this product is the same as that of $V(x_1, x_2, \ldots, x_n)$ and so $V(x_1, x_2, \ldots, x_n)$ is the product of all the $x_i - x_j$ (for $i < j$) times the product of all the x_i times a constant C.

Now considering the First Order Expansion, the coefficient of $x_1 x_2^2 \ldots x_n^n$ in $V(x_1, x_2, \ldots, x_n)$ is clearly 1 (coming from the diagonal components).

But the only term in $x_1 x_2^2 \ldots x_n^n$ in

$\prod_{i<j} (x_j - x_i)\, x_1 x_2 \ldots x_n$ comes from taking the x_j from each

$x_j - x_i$ and so its coefficient is 1**.**

Hence C = 1 and the theorem is proved.

Example 14: Evaluate the Vandermonde determinant

$$\begin{vmatrix} 2 & 4 & 6 & 16 \\ 3 & 9 & 27 & 81 \\ 4 & 16 & 64 & 256 \\ 5 & 25 & 125 & 625 \end{vmatrix}$$

Solution: $\begin{vmatrix} 2 & 4 & 6 & 16 \\ 3 & 9 & 27 & 81 \\ 4 & 16 & 64 & 256 \\ 5 & 25 & 125 & 625 \end{vmatrix}$

$= (5 - 4)(5 - 3)(5 - 2)(4 - 3)(4 - 2)(3 - 2)$

$= 1.2.3.1.2.1 = 12.$

§5.5. Elementary Operations and Elementary Matrices

We've spent some time talking about evaluating determinants in the most efficient manner. But who wants to evaluate a determinant? True, 3×3 determinants have a geometric significance, but what about a 10×10 determinant?

The answer is that the formula for the multiplicative inverse of an $n \times n$ matrix involves its determinant (as we saw for $n = 2$ and 3). And, moreover, the determinant *determines* whether or not an inverse exists. Zero determinant means no inverse. Non-zero determinant means there is an inverse. In order to prove this, and to derive the formula for the inverse, we need to discuss elementary row and column operations. These operations are not only a useful tool in proving things. They're a fundamental tool for doing lots of useful things with matrices – and they're especially relevant to the systematic solution of systems of linear equations. This is an extremely important application of matrices and we will devote the whole of the next chapter to this application.

There are five types of operations on the rows of a matrix that are particularly useful. They are called the **elementary row operations**. We denote the *i*'th row by $\mathbf{R_i}$.

ELEMENTARY ROW OPERATIONS	
$R_i \leftrightarrow R_j$	swap rows i, j (where $i \neq j$)
$R_i \times k$	multiply row i by k (where $k \neq 0$)
$R_i \div k$	divide row i by k (where $k \neq 0$)
$R_i - kR_j$	subtract k times row j from row i (where $i \neq j$)
$R_i + kR_j$	add k times row j to row i (where $i \neq j$)

NOTES:

(1) In fact there are really only three different types since

$R_i \div k$ is the same as $R_i \times (1/k)$ and

$R_i - kR_j$ is the same as $R_i + (-k)R_j$.

(2) With $R_i - kR_j$ and $R_i + kR_j$ only R_i changes.
Row j is not multiplied by k.

(3) $R_i + R_j$ looks symmetric in i, j but in fact it is R_j that gets added to R_i not the other way round.

(4) Each of the elementary row operations is reversible, and the inverse operation is also an elementary row operation. For example to reverse the operation of multiplying a row by k you simply divide it by k. And to reverse $R_i - kR_j$ you perform $R_i + kR_j$.

(5) $R_i \times 0$ is not an elementary row operation because it's not reversible.

(6) The row numbering refers to the current position, not the original position.

(7) We can rearrange the rows in any way by a sequence of row swaps. For example, to take the top row of a $m \times n$ matrix to the bottom and move every other row up you can perform:

$$R_1 \leftrightarrow R_2,\ R_2 \leftrightarrow R_3,\ \ldots\ R_{m-1} \leftrightarrow R_m.$$

For example if the matrix is $\begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{pmatrix}$, where the R_i represent four rows in terms of their original positions, the matrix changes as follows when the operations

$$R_1 \leftrightarrow R_2,\ R_2 \leftrightarrow R_3,\ R_3 \leftrightarrow R_4$$

are carried out:

$$\begin{pmatrix} R_1 \\ R_2 \\ R_3 \\ R_4 \end{pmatrix} \rightarrow \begin{pmatrix} R_2 \\ R_1 \\ R_3 \\ R_4 \end{pmatrix} \rightarrow \begin{pmatrix} R_2 \\ R_3 \\ R_1 \\ R_4 \end{pmatrix} \rightarrow \begin{pmatrix} R_2 \\ R_3 \\ R_4 \\ R_1 \end{pmatrix}.$$

ELEMENTARY COLUMN OPERATIONS	
$C_i \leftrightarrow C_j$	swap columns i, j (where $i \neq j$)
$C_i \times k$	multiply column i by k (where $k \neq 0$)
$C_i \div k$	divide column i by k (where $k \neq 0$)
$C_i - kC_j$	subtract k times column j from column i (where $i \neq j$)
$C_i + kC_j$	add k times row j to row i (where $i \neq j$)

The remarks about the elementary row operations apply to the elementary column operations, with this additional comment. You cannot mix rows and columns in the one operation. For example $R_i \leftrightarrow C_j$ and $R_i - kC_j$ are not elementary operations.

We've seen some of these ten operations in the context of determinants. Carrying out an elementary row or column operation on a square matrix changes its determinant in a very simple way.

Elementary operation	**Effect on the determinant**
$R_i \leftrightarrow R_j$ or $C_i \leftrightarrow C_j$	changes the sign
$R_i \times k$ or $C_i \times k$	multiplies the determinant by k
$R_i \div k$ or $C_i \div k$	divides the determinant by k
$R_i + kR_j$ or $C_i + kR_j$	no change
$R_i - kR_j$ or $C_i - kR_j$	no change

Example 15: Use elementary row operations to simplify the determinant

$$\begin{vmatrix} 1000 & 2000 & 3000 & 4000 \\ 2001 & 40002 & 6003 & 8004 \\ 3005 & 6006 & 9007 & 12008 \\ 4009 & 8010 & 12011 & 16012 \end{vmatrix}.$$

Solution: The answer is

$$\begin{vmatrix} 1000 & 2000 & 3000 & 4000 \\ 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{vmatrix} \quad (R_2 - 2R_1,\ R_3 - 3_1,\ R_4 - 4R_1$$

$$= 1000 \begin{vmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{vmatrix} \quad (R_1 \div 1000)$$

$$= 1000 \begin{vmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{vmatrix} \quad (R_1 - R_2)$$

$= 0.$

Each of the ten elementary row and column operations can be achieved by multiplying by a suitable matrix. For the row operations we must *pre-multiply*, that is, multiply on the left. For column operations we must *post-multiply*, that is, multiply on the right. The matrices that do the job are called **elementary matrices**. We denote the elementary matrix that corresponds to a given operation by enclosing the symbol for that operation in square brackets.

The elementary matrices are as follows:
$[R_i \leftrightarrow R_j] = [C_i \leftrightarrow C_j]$;
$[R_i + kR_j] = [C_j + \mathrm{k}C_i]$;
$[R_i - kR_j] = [C_j - kC_i]$;
$[R_i \times k] = [C_i \times k]$ (for $k \neq 0$).
$[R_i \div k] = [C_\mathrm{i} \div k]$ (for $k \neq 0$).

This isn't to say that $R_i \leftrightarrow R_j$ is the same operation as $C_i \leftrightarrow C_j$. However the same matrix achieves the row swap if it comes before the matrix and swaps the columns if it comes after. Notice the subtle change in the second case. Pre-multiplying by $[R_i - kR_j]$ changes row i but post-multiplying changes row j.

Now these elementary matrices are easy to write down. We simply carry out the operation on the identity matrix and we have the corresponding elementary matrix.

Theorem 12: The determinants of the elementary matrices are as follows:

Elementary Matrix	**Determinant**
$R_i \leftrightarrow R_j$	-1
$R_i \times k$	k
$R_i \div k$	$1/k$
$R_i + kR_j$	1
$R_i - kR_j$	1

Proof: Simply apply theorems 8-10 to the identity matrix.

Example 16: Find the elementary 3×3 matrices $[R_1 \leftrightarrow R_2]$, $[R_3 - kR_1]$ and $[R_2 \times k]$.

Solution: $[R_1 \leftrightarrow R_2] = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $[R_3 - kR_1] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -k & 0 & 1 \end{pmatrix}$

and $[R_2 \times k] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Note that $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ and

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{12} & a_{11} & a_{13} \\ a_{22} & a_{21} & a_{23} \\ a_{32} & a_{31} & a_{33} \end{pmatrix}.$$

Also $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -k & 0 & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$= \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31}-ka_{11} & a_{32}-ka_{12} & a_{33}-ka_{13} \end{pmatrix}$ and

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\lambda & 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11}-ka_{13} & a_{12} & a_{13} \\ a_{21}-ka_{23} & a_{22} & a_{23} \\ a_{31}-ka_{33} & a_{32} & a_{33} \end{pmatrix}.$$

An important family of $n \times n$ matrices are the ‘cousins’ of the identity matrix.

For $0 \leq r \leq n$ we define I_r to be the diagonal matrix with the first r components on the diagonal being 1 and the rest being zero. So I_n is, in fact, the identity matrix I. Also I_0 is the zero matrix.

Example 17: The 4×4 matrix $I_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

Theorem 13: Multiplying an $n \times n$ matrix on the left by I_r (for $r < n$) makes its determinant zero.
Proof: We prove this by induction on n using the 1st order expansion.

§5.6. Matrix Factorisation

Every positive integer can be factorised uniquely into primes. Something similar holds for square matrices, except that the factorisation is far from unique. Mostly these factors can be elementary matrices but as these are all invertible and some matrices are not, we need to supplement them by the I_r matrices.

A **prime** matrix is one that is either an elementary matrix or an I_r matrix.

Theorem 14: Every $n \times n$ matrix A can be converted into an I_r matrix by a sequence of elementary row and column operations.
Proof: We prove this by induction on n.
If $n = 1$ then $A = (a)$ for some a.
If $a = 0$ then $A = I_0$.
If $a \neq 0$ then dividing by a converts A to $(1) = I_1$.

Suppose that A is $n \times n$ where $n > 1$ and suppose that the theorem holds for smaller matrices.
If $A = 0$ then $A = I_0$.

Suppose that A has a non-zero component, a. By swapping rows and swapping columns this can be brought into the 1-1 position.

By dividing the first row by a this 1-1 component becomes 1. We then subtract suitable multiples of the first row to get 0's underneath this 1. By subtracting suitable multiples of the first column we can get 0's to the right of the 1.

So, by a sequence of elementary row and column operations A can be put in the form $\begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & B \end{pmatrix}$ where B is an $(n-1) \times (n-1)$ matrix.

By induction B can be converted to the I_r form by a sequence of elementary row and column operations. Although these only operate on B the same operations applied to the whole n rows and n columns will have an identical effect because the extra components are 0's. So the matrix can be put in the form $\begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & I_r \end{pmatrix} = I_{r+1}$ by a sequence of elementary operations.

Corollary: Every square matrix, A, is a product of prime matrices.

Proof: Every elementary row operation is equivalent to pre-multiplying by an elementary matrix and every elementary column operation is equivalent to post-multiplying by an elementary matrix.

Hence $A = EI_rF$ where E and F are products of elementary matrices.

NOTE: Unlike many factorisations into primes in mathematics this factorisation is not unique.

Example 18: Factorise $\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{pmatrix}$ into prime matrices.

Solution: We already have a non-zero component in the 1-1 position and we don't have to divide to make it 1.

So now apply the following sequence of elementary row operations: $R_2 - 4R_1$ and $R_3 - 7R_1$.

This gives us $\begin{pmatrix}1 & 2 & 3\\ 0 & -3 & -6\\ 0 & -6 & -12\end{pmatrix}$.

Hence $[R_3 - 7R_1]\,[R_2 - 4R_1]\,A = \begin{pmatrix}1 & 2 & 3\\ 0 & -3 & -6\\ 0 & -6 & -12\end{pmatrix}$.

Note that we have reversed the order because we are multiplying on the left, though since these operations are independent and could be carried out in either order, it doesn't matter.

Now we carry out $C_2 - 2C_1$ followed by $C_3 - 3C_1$. This merely wipes out the rest of the first row. Thus:

$$[R_3 - 7R_1]\,[R_2 - 4R_1]\,A\,[C_2 - 2C_1]\,[C_3 - 3C_1]$$

$$= \begin{pmatrix}1 & 0 & 0\\ 0 & -3 & -6\\ 0 & -6 & -12\end{pmatrix}.$$

Divide the second row by -3. This gives:

$[R_2 \div (-3)]\,[R_3 - 7R_1]\,[R_2 - 4R_1]\,A\,[C_2 - 2C_1]\,[C_3 - 3C_1]$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & -6 & -12 \end{pmatrix}.$$

Add 6 times row 2 to row 3. Hence:

$[R_3 + 6R_2]\ [R_2 \div (-3)]\ [R_3 - 7R_1]\ [R_2 - 4R_1]\ A$

$$\times [C_2 - 2C_1] \times [C_3 - 3C_1] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Finally we subtract 2 times column 2 from column 3.

$[R_3 + 6R_2]\ [R_2 \div (-3)]\ [R_3 - 7R_1]\ [R_2 - 4R_1]\ A$

$$\times [C_2 - 2C_1]\ [C_3 - 3C_1]\ [C_3 - 2C_2] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = I_2.$$

Hence

$$A = [R_2 - 4R_1]^{-1}\ [R_3 - 7R_1]^{-1}[R_2 \div (-3)]^{-1}[R_3 + 6R_2]^{-1}\ I_2$$
$$\times [C_3 - 2C_2]^{-1}[C_3 - 3C_1]^{-1}[C_2 - 2C_1]^{-1}$$
$$= [R_1 + 4R_2]\ [R_1 + 7R_3]\ [R_2 \times (-3)]\ [R_2 - 6R_3]\ I_2$$
$$\times [C_2 + 2C_3]\ [C_1 + 3C_3]\ [C_1 + 2C_2]$$

which is a product of prime matrices.

NOTE: Remember that $[R_i - kR_j]^{-1} = [R_j + kR_i]$ etc.

Theorem 15: If P is a prime $n \times n$ matrix then $|PB| = |P|.|B|$ for all $n \times n$ matrices B.

Proof: If $P = [R_i \leftrightarrow R_j]$ then

$|PB| = -|B| = (-1)|B| = |[R_i \leftrightarrow R_j]|.|B| = |P|.|B|$.

Similarly if $P = [R_i + kR_j]$, $[R_i - kR_j]$, $[R_i \times k]$ or $[R_i \div k]$.

If $P = I_n = I$ then $|PB| = |B| = |P|.|B|$.

If $P = I_r$ for $r < n$ then PB has at least one row of zeros at the bottom and so $|PB| = 0 = |P|.|B|$.

Theorem 16: For all $n \times n$ matrices, $|AB| = |A|.|B|$.
Proof: Let $A = P_1P_2 \ldots P_k$ where each P_i is a prime matrix.
Then $|AB| = |P_1P_2 \ldots P_kB|$

$$= |P_1|.|P_2|. \ldots |P_k|.|B|$$
$$= |P_1P_2 \ldots P_k|.|B|$$
$$= |A|.|B|.$$

Theorem 17: If P is a elementary matrix then $|P^T| = |P|$.
Proof: $[R_i \leftrightarrow R_j]$, $[R_i \times \lambda]$, $[R_i \div \lambda]$ and I_r are symmetric.
$[R_i \pm k\, R_j]^T = [R_j \pm kR_i]$.

Theorem 18: $|A^T| = |A|$.
Proof: Let $A = P_1P_2 \ldots P_k$ where the P_i are prime matrices.
Then $|A^T| = |P_k^T|. \ldots |P_2^T|.|P_1^T|$

$$= |P_k|. \ldots |P_2|.|P_1|$$
$$= |P_1|.|P_2|. \ldots |P_k|$$
$$= |A|.$$

Corollary: Theorems 8, 9 and 10 hold if 'row' is replaced by 'column'.

Theorem 19 (Expansion along any row):

For any i, $|A| = \sum_k (-1)^{i+k} a_{ik}\, |\delta_k^i(A)|$.

Proof: If $i = 1$ this is just the definition.
Suppose i > 1.
Let B = $[R_1 \leftrightarrow R_i]$A. Expanding |B| along the i'th row, |B|
$$= -\sum_k (-1)^{i+k} a_{ik} |\delta_k^i(A)| = -|A|.$$
Because of this theorem you can expand along *any* row, by multiplying each entry by the determinant obtained by deleting that row and column. These products are added and subtracted. The signs alternate. Whether the first sign is + or – depends on which row you are using. The signs follow a chessboard pattern with + in the top left hand corner:

$$\begin{matrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \dots & & & & \end{matrix}$$

Similarly one can expand down any column in a similar way, since transposing leaves a determinant unchanged.

Example 19: Evaluate $\begin{vmatrix} 12 & 4 & 6 \\ 10 & 0 & 83 \\ 0 & 0 & 49 \end{vmatrix}$.

Solution: Expanding along the 3rd row we get
$$49\begin{vmatrix} 12 & 4 \\ 10 & 0 \end{vmatrix} = (49)(-40) = -1960.$$
Expanding down the 2nd column we get

$(-4)\begin{vmatrix}10 & 83\\ 0 & 49\end{vmatrix} = (-4)(490) = -1960.$

§5.7. Inverse of a Matrix

By theorem 14 every square matrix can be put in the form $A = E_h \ldots E_1 I_r F_1 \ldots F_k$.

If $r = n$ we can write A as a product of elementary matrices so every square matrix is covered by either Theorem 20 or Theorem 21 as follows.

Theorem 20: Suppose $A = P_1P_2 \ldots P_k$ where each P_i is an elementary matrix.
Then:

(1) $|A| \neq 0$;
(2) A is invertible;
(3) $A\mathbf{v} = \mathbf{0}$ implies that $\mathbf{v} = \mathbf{0}$.

Proof: (1) $|A| = |P_1|.|P_2| \ldots |P_k|$ and $|P_i| \neq 0$ for each i.
(2) $A^{-1} = P_k^{-1} \ldots P_2^{-1}P_1^{-1}$.
(3) If $A\mathbf{v} = \mathbf{0}$ then $\mathbf{v} = A^{-1}\mathbf{0} = \mathbf{0}$.

Theorem 21: Suppose the $n \times n$ matrix $A = P_1P_2 \ldots P_k$ where each P_i is a prime matrix and at least one $P_i = I_r$ for some $r < n$.
Then:

(1) $|A| = 0$;
(2) there exists $\mathbf{v} \neq 0$ such that $\mathbf{Av} = \mathbf{0}$;
(3) A is not invertible.

Proof:

(1) Since $r < n$, $|I_r| = 0$ and hence:

$$|A| = |P_1|.|P_2| \dots |P_k| = 0.$$

(2) Let m be the largest value of i such that P_i is an I_r matrix for $r \le k$.

So any factor to the right of P_m is an elementary matrix and so is invertible.

Let $\mathbf{v} = P_k^{-1} \dots P_{s+1}^{-1}\begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \end{pmatrix}$.

Clearly $\mathbf{v} \neq 0$. Note that if $s = k$ then $\mathbf{v} = \begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \end{pmatrix}$.

Then $A\mathbf{v} = P_1P_2 \dots P_{m-1}I_rP_{s+1} \dots P_k\, P_k^{-1} \dots P_{m+1}^{-1}\begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \end{pmatrix}$

$$= P_1P_2 \dots P_{m-1}I_r\begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \end{pmatrix}$$

$$= \mathbf{0} \text{ since } I_r\begin{pmatrix} 0 \\ \dots \\ 0 \\ 1 \end{pmatrix} = \mathbf{0}.$$

(3) Suppose A^{-1} exists and let m be the smallest value of I such that P_i is an I_r matrix for $r \leq k$.
So any factor to the left of P_m is an elementary matrix and so is invertible.

Then $P_{m-1}^{-1} \ldots P_1^{-1} A = I_r P_{m+1} \ldots P_k$

Hence $I_r(P_{m+1} \ldots P_k A^{-1} P_1 \ldots P_{m-1})$

$$= (I_r P_{m+1} \ldots P_k)(A^{-1} P_1 \ldots P_{m-1})$$
$$= (P_{m-1}^{-1} \ldots P_1^{-1} A)(A^{-1} P_1 \ldots P_{m-1})$$
$$= I.$$

But the last row of the left hand side is zero, a contradiction!

We'd now like to be able to write down a formula for the inverse of an invertible square matrix. One method involves the use of cofactors.

If $A = (a_{ij})$ is an $n \times n$ matrix we define the i-j **cofactor** to be $A_{ij} = (-1)^{i+j}|\delta^i_j(A)|$.

That is, we delete the i'th row and j'th column and evaluate the $(n-1) \times (n-1)$ determinant of what remains. Then, if $i + j$ is odd, we change the sign.

The **adjoint** of **A** is defined to be **adj(A)** $= (A_{ij})^T$, that is, the matrix of cofactors transposed. This method is known as the **adjoint method** or **cofactor method**.

It's clear that, unless n is quite small, this involves an enormous amount of computation. For a 100×100 matrix we need to evaluate ten thousand 99×99 determinants, each of which involves evaluating ninety-nine 98×98 determinants, and so on. The adjoint can be used to find inverses, but computationally there's a much more efficient method that we'll discuss later. However the adjoint method does have some theoretical interest.

Theorem 22: A.adj(A) = |A|.I = adj(A).A

Proof: The i-j component of A.adj(A) is $\sum_k a_{ik} A_{kj}$.

The i-i component of this product is

$$\sum_k a_{ik} (-1)^{i+k} |\delta_k^i(A)| = |A|.$$

If $i \neq j$ the i-j component is $\sum_k a_{ik} (-1)^{j+k} |\delta_k^j(A)|$.

This is independent of the j'th row. So, if we replace the j'th row by a second copy of the i'th to obtain $B = (b_{ij})$ then

$$\sum_k a_{ik} (-1)^{j+k} |\delta_k^j(A)| = \sum_k a_{jk} (-1)^{j+k} |\delta_k^j(B)|$$

$$= |B| = 0,$$

since B has two identical rows.

Corollary: If A is invertible $A^{-1} = \frac{1}{|A|}$ adj(A).

Example 20: Find the inverse of $A = \begin{pmatrix} 4 & 8 & 9 \\ 5 & 2 & 7 \\ 2 & 13 & 23 \end{pmatrix}$.

$$\text{adj}A = \begin{pmatrix} \begin{vmatrix} 2 & 7 \\ 13 & 23 \end{vmatrix} & -\begin{vmatrix} 5 & 7 \\ 2 & 23 \end{vmatrix} & \begin{vmatrix} 5 & 2 \\ 2 & 13 \end{vmatrix} \\ -\begin{vmatrix} 8 & 9 \\ 13 & 23 \end{vmatrix} & \begin{vmatrix} 4 & 9 \\ 2 & 23 \end{vmatrix} & -\begin{vmatrix} 4 & 8 \\ 2 & 13 \end{vmatrix} \\ \begin{vmatrix} 8 & 9 \\ 2 & 7 \end{vmatrix} & -\begin{vmatrix} 4 & 9 \\ 5 & 7 \end{vmatrix} & \begin{vmatrix} 4 & 8 \\ 5 & 2 \end{vmatrix} \end{pmatrix}^T$$

$$= \begin{pmatrix} 46 - 91 & 14 - 115 & 65 - 4 \\ 117 - 184 & 92 - 18 & 52 - 16 \\ 56 - 18 & 45 - 28 & 8 - 40 \end{pmatrix}$$

$$= \begin{pmatrix} -45 & -101 & 61 \\ -67 & 74 & 36 \\ 38 & 17 & -32 \end{pmatrix}^T$$

$$= \begin{pmatrix} -45 & -67 & 38 \\ -101 & 74 & 17 \\ 61 & 36 & -32 \end{pmatrix}.$$

$$|A| = (4, 8, 9)\begin{pmatrix} -45 \\ -101 \\ 61 \end{pmatrix} = -180 - 808 + 549 = -439.$$

Hence $A^{-1} = \begin{pmatrix} 45/439 & 67/439 & -38/439 \\ 101/439 & -74/439 & -17/439 \\ -61/439 & -36/439 & 32/439 \end{pmatrix}$

As I said earlier, this nice neat formula is not efficient computationally. The following is the one that's normally used in practice. If A is an $n \times n$ matrix we adjoin the $n \times n$ identity matrix to the right of A and so consider the $n \times 2n$ matrix (A | I). We now transform A to the identity matrix using only elementary row operations. Of course, if A is not invertible this won't be possible and we'll get a row of zeros in the left half. This will alert us to the fact that A^{-1} doesn't exist.

Every elementary row operation is carried out on both halves of the array, so that whatever is done to A is done to I. Once I appears in the left half of the array, the inverse A^{-1} will appear in the right half.

The reason why this works is very simple. If E_1, E_2, … , E_k are the matrices of the elementary row operations that transform A to I then:

$$E_k \ldots E_2E_1A = I, \text{ so that:}$$

$$E_k \ldots E_2E_1 = A^{-1}.$$

If B is the matrix that appears in the right half of the array at the end then $E_k \ldots E_2E_1I = B$ and so $B = A^{-1}$.

$$\mathbf{(A \mid I) \rightarrow (I \mid A^{-1})}$$

Example 21: Find the inverse of the matrix $A = \begin{pmatrix} 3 & 4 & 1 \\ 0 & 5 & -1 \\ 2 & 3 & 7 \end{pmatrix}$

by:

(1) computing the adjoint;

(2) by using the Gauss-Jordan reduction whereby (A|I) is transformed to $(I|A^{-1})$.

$$\textbf{Solution: } (1)\ \text{adj}A = \begin{pmatrix} \begin{vmatrix} 5 & -1 \\ 3 & 7 \end{vmatrix} & -\begin{vmatrix} 0 & -1 \\ 2 & 7 \end{vmatrix} & \begin{vmatrix} 0 & 5 \\ 2 & 3 \end{vmatrix} \\ -\begin{vmatrix} 4 & 1 \\ 3 & 7 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 2 & 7 \end{vmatrix} & -\begin{vmatrix} 3 & 4 \\ 2 & 3 \end{vmatrix} \\ \begin{vmatrix} 4 & 1 \\ 5 & -1 \end{vmatrix} & -\begin{vmatrix} 3 & 1 \\ 0 & -1 \end{vmatrix} & \begin{vmatrix} 3 & 4 \\ 0 & 5 \end{vmatrix} \end{pmatrix}^T$$

$$= \begin{pmatrix} 38 & -2 & -10 \\ -25 & 19 & -1 \\ -9 & 3 & 15 \end{pmatrix}^T = \begin{pmatrix} 38 & -25 & -9 \\ -2 & 19 & 3 \\ -10 & -1 & 15 \end{pmatrix}$$

$$|A| = 3\begin{vmatrix} 5 & -1 \\ 3 & 7 \end{vmatrix} - 4\begin{vmatrix} 0 & -1 \\ 2 & 7 \end{vmatrix} + \begin{vmatrix} 0 & 5 \\ 2 & 3 \end{vmatrix}$$

$$= 3(38) - 4(2) - 10$$

$$= 114 - 8 - 10$$

$$= 96.$$

So $A^{-1} = \dfrac{1}{96}\begin{pmatrix} 38 & -25 & -9 \\ -2 & 19 & 3 \\ -10 & -1 & 15 \end{pmatrix}$.

(2) $$\left(\begin{array}{ccc|ccc} 3 & 4 & 1 & 1 & 0 & 0 \\ 0 & 5 & -1 & 0 & 1 & 0 \\ 2 & 3 & 7 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 1 & -6 & 1 & 0 & -1 \\ 0 & 5 & -1 & 0 & 1 & 0 \\ 2 & 3 & 7 & 0 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 1 & -6 & 1 & 0 & -1 \\ 0 & 5 & -1 & 0 & 1 & 0 \\ 0 & 1 & 19 & -2 & 0 & 3 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 1 & -6 & 1 & 0 & -1 \\ 0 & 1 & 19 & -2 & 0 & 3 \\ 0 & 5 & -1 & 0 & 1 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 1 & -6 & 1 & 0 & -1 \\ 0 & 1 & 19 & -2 & 0 & 3 \\ 0 & 0 & -96 & 10 & 1 & -15 \end{array}\right)$$

$$\rightarrow\left(\begin{array}{ccc|ccc} 1 & 1 & -6 & 1 & 0 & -1 \\ 0 & 1 & 19 & -2 & 0 & 3 \\ 0 & 0 & 1 & -10/96 & -1/96 & 15/96 \end{array}\right)$$

$$\rightarrow\left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 36/96 & -6/96 & -6/96 \\ 0 & 1 & 0 & -2/96 & 19/96 & 3/96 \\ 0 & 0 & 1 & -10/96 & -1/96 & 15/96 \end{array}\right)$$

$$\rightarrow\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 38/96 & -25/96 & -9/96 \\ 0 & 1 & 0 & -2/96 & 19/96 & 3/96 \\ 0 & 0 & 1 & -10/96 & -1/96 & 15/96 \end{array}\right) \text{ so}$$

$$A^{-1} = \frac{1}{96}\begin{pmatrix} 38 & -25 & -9 \\ -2 & 19 & 3 \\ -10 & -1 & 15 \end{pmatrix}.$$

You'd be correct if you felt that the adjoint method involved less computation, and so you should use the adjoint method for finding the inverses of 2×2 and 3×3 matrices. But for 4×4 matrices and larger there's less work if you use the (A|I) method, and the larger the matrix the more dramatic is the difference. A computer would be hard put to invert a 100×100 matrix (such large matrices do arise in certain engineering applications) but it could cope easily doing it by the (A|I) method.

§5.8. Upper and Lower Triangular Matrices

The **diagonal** of a matrix is the one that goes from the top left. Its components are of the form a_{ii}. If the matrix is square this diagonal will reach the bottom right hand corner. We'll consider here only square matrices.

A **diagonal matrix** is a square matrix where every component above or below the diagonal is zero. So A = (a_{ij}) is diagonal if $a_{ij} = 0$ whenever $i \neq j$. Some or all of the diagonal components can also be zero. Special examples are the zero matrix and the identity matrix.

So a diagonal matrix has the form

$$\begin{pmatrix} d_1 & 0 & \ldots\ldots & 0 \\ 0 & d_2 & \ldots\ldots & 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots & \ldots\ldots \\ 0 & 0 & \ldots\ldots & d_n \end{pmatrix}.$$

Sometimes we write this as **diag(d_1, d_2, ..., d_n)**.

Clearly the sum, difference and product of two $n \times n$ diagonal matrices is again diagonal.

Example 22: If A = diag($a_1, a_2, \ldots, a_k$) and

B = diag($b_1, b_2, \ldots, b_k$)

then A + B = diag($a_1 + b_1, a_2 + b_2, \ldots, a_k + b_k$),

A – B = diag($a_1 - b_1, a_2 - b_2, \ldots, a_k - b_k$) and

AB = diag($a_1b_1, a_2b_2, \ldots, a_kb_k$).

In particular, if m is a positive integer, the m'th power of A is = diag($a_1^n, a_2^n, \ldots, a_k^n$).

If each $a_i \neq 0$ then A is invertible and

$$A^{-1} = \mathrm{diag}(a_1^{-1}, a_2^{-1}, \ldots, a_k^{-1}).$$

An $n \times n$ matrix A = (a_{ij}) is **upper-triangular** if $a_{ij} = 0$ when $i > j$. This means that the components below the diagonal are zero.

The matrix A is **lower-triangular** if $a_{ij} = 0$ when $i < j$, that is, the components above the diagonal are zero.

Example 23: A 4 × 4 upper-triangular matrix has the form

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{pmatrix}$$

and a 4 × 4 lower-triangular matrix has the form

$$\begin{pmatrix} a_{11} & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ a_{31} & a_{32} & a_{33} & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}.$$

It's easy to see, by expanding along the first row (for lower-triangular matrices) or down the first column (for upper-triangular matrices) that the determinant of a triangular matrix (upper or lower) is simply the product of the diagonal components.

If A and B are $n \times n$ upper-triangular matrices then so are A + B and A – B. It's not difficult to see that AB is also upper-triangular.
It'is also true that if the diagonal components of an upper-triangular matrix A are non-zero then A is invertible and A^{-1} is upper-triangular. The easiest way to prove this is by induction, using partitioned matrices. Similar results hold for lower-triangular matrices.

If A is an $n \times n$ matrix and $0 < r < n$, then we may represent A in the form $\begin{pmatrix} B & C \\ D & E \end{pmatrix}$ where B is the $r \times r$ matrix consisting of the components in the first r rows and columns. C, D, E are defined similarly. If $s = n - r$ then C is $r \times s$, D is $s \times r$ and E is $s \times s$.

We say that such a matrix has been **partitioned**.
If two $n \times n$ matrices $\begin{pmatrix} B_1 & C_1 \\ D_1 & E_1 \end{pmatrix}$ and $\begin{pmatrix} B_2 & C_2 \\ D_2 & E_2 \end{pmatrix}$ are partitioned in the same way (that is, with the same r, s) then:

$$\begin{pmatrix} B_1 & C_1 \\ D_1 & E_1 \end{pmatrix} + \begin{pmatrix} B_2 & C_2 \\ D_2 & E_2 \end{pmatrix} = \begin{pmatrix} B_1 + B_2 & C_1 + C_2 \\ D_1 + D_2 & E_1 + E_2 \end{pmatrix} \text{ and}$$

$$\begin{pmatrix} B_1 & C_1 \\ D_1 & E_1 \end{pmatrix} + \begin{pmatrix} B_2 & C_2 \\ D_2 & E_2 \end{pmatrix} = \begin{pmatrix} B_1B_2 + C_1D_2 & B_1C_2 + C_1E_2 \\ D_1B_2 + E_1D_2 & D_1C_2 + E_1E_2 \end{pmatrix}$$

In other words you multiply partitioned matrices as if the matrix components are just ordinary numbers. But you should assure yourself that all the individual products are defined and that the overall product is partitioned in the same way as the partitioned factors.

Suppose B_1 and B_2 are $r \times r$, and C_1 is $r \times (n - r)$ while D_2 is $(n - r) \times r$. Thus C_1D_2 is defined and is $r \times r$ and so can be added to B_1B_2.

Theorem 24: If $A = \begin{pmatrix} B & C \\ 0 & E \end{pmatrix}$ is an invertible partitioned upper-triangular matrix then $A^{-1} = \begin{pmatrix} B^{-1} & -B^{-1}CE^{-1} \\ 0 & E^{-1} \end{pmatrix}$.

Proof: $$\begin{pmatrix} B & C \\ 0 & E \end{pmatrix}\begin{pmatrix} B^{-1} & -B^{-1}CE^{-1} \\ 0 & E^{-1} \end{pmatrix} = \begin{pmatrix} I & -CE^{-1} + CE^{-1} \\ 0 & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

§5.9. LU Factorisation

Sometimes we have to solve systems of the form $A\mathbf{v} = \mathbf{b}$ where we have the same A and many **b**'s. The steps in the Gaussian algorithm that reduce the matrix to echelon form would be the same in each case, resulting in lots of repeated calculation. One way to avoid this inefficiency is to put all the constant vectors into a single matrix and so solve all the systems at the same time.

Example 24: Solve the systems $\left.\begin{array}{r} 2x + 3y + 2z = 7 \\ 5x + y + 4z = 36 \\ 3x + 7y + z = 1 \end{array}\right\}$ and $\left.\begin{array}{r} 2x + 3y + 2z = -1 \\ 5x + y + 4z = -25 \\ 3x + 7y + z = 31 \end{array}\right\}$.

Solution: The augmented matrix is $\left(\begin{array}{ccc|cc} 2 & 3 & 2 & 7 & -1 \\ 5 & 1 & 4 & 36 & -25 \\ 3 & 7 & 1 & 1 & 31 \end{array}\right)$.

$$\left(\begin{array}{ccc|cc} 2 & 3 & 2 & 7 & -1 \\ 5 & 1 & 4 & 36 & -25 \\ 3 & 7 & 1 & 1 & 31 \end{array}\right) \to \left(\begin{array}{ccc|cc} 2 & 3 & 2 & 7 & -1 \\ 1 & -5 & 0 & 22 & -23 \\ 1 & 4 & -1 & -6 & 32 \end{array}\right)$$

$R_2 - 2R_1,\ R_3 - R_1$

$$\to \left(\begin{array}{ccc|cc} 1 & 4 & -1 & -6 & 32 \\ 2 & 3 & 2 & 7 & -1 \\ 1 & -5 & 0 & 22 & -23 \end{array}\right)$$

$R_1 \leftrightarrow R_3,\ R_2 \leftrightarrow R_3$

$$\to \left(\begin{array}{ccc|cc} 1 & 4 & -1 & -6 & 32 \\ 0 & -5 & 4 & 19 & -65 \\ 0 & -9 & 1 & 28 & -55 \end{array}\right)$$

$R_2 - 2R_1,\ R_3 - R_1$

$$\to \left(\begin{array}{ccc|cc} 1 & 4 & -1 & -6 & 32 \\ 0 & -5 & 4 & 19 & -65 \\ 0 & 1 & -7 & -10 & 75 \end{array}\right)$$

$R_3 - 2R_2$

$$\rightarrow \left(\begin{array}{ccc|cc} 1 & 4 & -1 & -6 & 32 \\ 0 & 1 & -7 & -10 & 75 \\ 0 & -5 & 4 & 19 & -65 \end{array}\right) \quad R_2 \leftrightarrow R_3$$

$$\rightarrow \left(\begin{array}{ccc|cc} 1 & 4 & -1 & -6 & 32 \\ 0 & 1 & -7 & -10 & 75 \\ 0 & 0 & -31 & -31 & 310 \end{array}\right) \quad R_3 + 5R_2$$

$$\rightarrow \left(\begin{array}{ccc|cc} 1 & 4 & -1 & -6 & 32 \\ 0 & 1 & -7 & -10 & 75 \\ 0 & 0 & 1 & 1 & -10 \end{array}\right) \quad R_3 \div (-30).$$

We can use back-substitution or continue till the left hand matrix is the identity. Either way we get the solution:

$x = 7$, $y = -3$, $z = 1$ for the first system and

$x = 2$, $y = 5$, $z = -10$ for the second.

Often, however, we don't know all the constant terms at the same time. The constant terms might arise from daily readings of some instrument. In this case, if the matrix of coefficients, A, was square, and invertible, we could find A^{-1} once and every day we'd simply calculate $A^{-1}\mathbf{b}$.

But, if A is not square, or perhaps square but not invertible, we'd have to go about it quite differently. In this case we factorise A as A = LU where L is a square and invertible lower triangular matrix and U is an upper triangular matrix. The method is known as the **LU method**, but in fact we'll be writing $A = E^{-1}U$, where $L = E^{-1}$, but the name '$E^{-1}U$ method' doesn't sound as good as the LU method!

Theorem 25: Every matrix A can be factorised as $A = E^{-1}U$ where E is invertible and lower triangular and U is an upper triangular matrix.
Solution: Suppose A is $m \times n$. Then adjoin the $m \times m$ identity matrix to the right of A and use elementary row operations until A is in echelon form. Let U be the left hand matrix and let E be the right hand matrix.

Being an echelon form, U will be an $m \times n$ upper-triangular matrix and E, being the product of the elementary matrices corresponding to the elementary row operations, will be an $m \times m$ invertible matrix.

Since the same set of elementary row operations transforms A into U as transforms I into E, we have $EA = U$. Hence $A = E^{-1}U$.

Corollary: Every matrix A can be factorised as A = LU where L is an invertible lower-triangular matrix and U is an upper triangular matrix.
Proof: Put $L = E^{-1}$. Clearly L is invertible. It remains to show that L is lower-triangular. This follows from the fact that the inverse of an invertible lower-triangular matrix is lower-triangular.

It should be pointed out that these factorisations are not unique. They depend on the particular sequence of elementary row operations that we choose.

But how does this help us to solve lots of systems of linear equations with the same matrix of coefficients? Suppose we want to solve A**v** = **b**, with lots of different **b**'s. We carry out an LU factorisation once, getting A = $E^{-1}U$ as above. Then the equation becomes $E^{-1}U$**v** = **b**, which gives U**v** = E**b**. So for each **b** we carry out matrix multiplication to find E**b** and then use back-substitution to solve U**v** = E**b**. The fact that U is upper-triangular makes this easy.

Example 25: Let $A = \begin{pmatrix} 3 & -1 & 2 & -4 & 1 \\ -3 & 3 & =5 & 5 & -2 \\ 6 & -4 & 11 & =10 & 6 \\ =6 & 8 & -21 & 13 & -9 \end{pmatrix}$.

(i) Write A in the $E^{-1}U$ form;

(ii) If **b** = $\begin{pmatrix} 1 \\ -2 \\ 9 \\ -15 \end{pmatrix}$, solve the system A**v** = **b**.

Solution:

(i) $$\left(\begin{array}{ccccc|cccc} 3 & -1 & 2 & -4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 3 & -5 & 5 & -2 & 0 & 1 & 0 & 0 \\ 6 & -4 & 11 & -10 & 6 & 0 & 0 & 1 & 0 \\ -6 & 8 & -21 & 13 & -9 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccccc|cccc} 3 & -1 & 2 & -4 & 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & -3 & 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & -2 & 7 & -2 & 4 & -2 & 0 & 1 & 0 \\ 0 & 6 & -17 & 5 & -7 & 2 & 0 & 0 & 1 \end{array}\right)$$

$R_2 + R_1$, $R_3 - 2R_1$, $R_4 + 2R_1$

$$\to \left(\begin{array}{ccccc|cccc} 3 & -1 & 2 & -4 & 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & -3 & 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 4 & -1 & 3 & -1 & 1 & 1 & 0 \\ 0 & 0 & -8 & 2 & -4 & -1 & -3 & 0 & 1 \end{array}\right)$$

$R_3 + R_2$, $R_4 - 3R_2$

$$\to \left(\begin{array}{ccccc|cccc} 3 & -1 & 2 & -4 & 1 & 1 & 0 & 0 & 0 \\ 0 & 2 & -3 & 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 4 & -1 & 3 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 & -3 & -1 & 2 & 1 \end{array}\right)$$

$R_4 + 2R_3$

So $U = \begin{pmatrix} 3 & -1 & 2 & -4 & 1 \\ 0 & 2 & -3 & 1 & -1 \\ 0 & 0 & 4 & -1 & 3 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$ and $E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -3 & -1 & 2 & 1 \end{pmatrix}$

.

(ii) $E\mathbf{b} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ -1 & 1 & 1 & 0 \\ -3 & -1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 9 \\ -15 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 6 \\ 2 \end{pmatrix}$.

$(U \mid E\mathbf{b}) = \left(\begin{array}{ccccc|c} 3 & -1 & 2 & -4 & 1 & 1 \\ 0 & 2 & -3 & 1 & -1 & -1 \\ 0 & 0 & 4 & -1 & 3 & 6 \\ 0 & 0 & 0 & 0 & 2 & 2 \end{array}\right)$.

By back-substitution we get

$x_5 = 1$, $x_4 = k$, $4x_3 = 6 + k - 3 = 3 + k$, so $x_3 = \dfrac{3+k}{4}$,

$2x_2 = -1 + 3\left(\dfrac{3+k}{4}\right) - k + 1 = \dfrac{9-k}{4}$, so $x_2 = \dfrac{9-k}{8}$,

$3x_1 = 1 + \left(\dfrac{9-k}{8}\right) - 2\left(\dfrac{3+k}{4}\right) + 4k - 1 = \dfrac{-3+27k}{8}$,

so $x_1 = \dfrac{9k-1}{8}$.

We can summarise the $E^{-1}U$ method as follows:

$(A \mid I) \rightarrow (U \mid E)$
To solve Ax = b, solve Ux = Eb.

EXERCISES FOR CHAPTER 5

Exercise 1: Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 & 9 & 8 \\ 2 & 4 & 6 & 8 \\ 5 & 3 & 1 & 0 \end{pmatrix}$, $C = \begin{pmatrix} 6 & 2 & 5 \end{pmatrix}$,
$D = \begin{pmatrix} 8 & 7 \\ 6 & 5 \end{pmatrix}$.

Find, where they exist: (i) $A + A^T$; (ii) $D - D^T$;
(iii) 2B; (iv) AD; (v) AB;
(vi) BA; (vii) AA^T; (viii) A^TA;
(ix) CBA; (x) $(A^TA)^2 - 1000D$.

Exercise 2: Find $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{vmatrix}$ by:

(i) the First Order Expansion;
(ii) the Second Order Expansion;
(iii) simplifying using elementary row and column operations.

Exercise 3: Find the inverse of A = $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$ by

(a) the adjoint method;
(b) the (A|I) method.
Check your answer by multiplying A by your inverse.

Exercise 4: Find the inverse of $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 5 & 6 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Check your answer.

Exercise 5:
(i) Prove that if A = BB^T for some matrix B then A is symmetric.
(ii) Prove that if A = $\begin{pmatrix} a & b \\ b & d \end{pmatrix}$ is a 2×2 real matrix then
A = BB^T for some real matrix B if and only if

$$a \geq 0, \; b \geq 0 \text{ and } ad - b^2 \geq 0.$$

Exercise 6:

Let A = $\left(\begin{array}{ccccc|cccc} 2 & -1 & 2 & -1 & 3 & 1 & 0 & 0 & 0 \\ 4 & 1 & 6 & -5 & 8 & 0 & 1 & 0 & 0 \\ 2 & -10 & -4 & 8 & -5 & 0 & 0 & 1 & 0 \\ 2 & -13 & -6 & 16 & -5 & 0 & 0 & 0 & 1 \end{array}\right)$.

Write A in the $E^{-1}U$ form and hence solve $A\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$.

SOLUTIONS FOR CHAPTER 5

Exercise 1: (i) undefined; (ii) $\begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$;

(iii) $\begin{pmatrix} 0 & 2 & 18 & 16 \\ 4 & 12 & 6 & 16 \\ 10 & 6 & 2 & 0 \end{pmatrix}$; (iv) $\begin{pmatrix} 20 & 17 \\ 48 & 41 \\ 76 & 65 \\ 104 & 89 \end{pmatrix}$; (v) undefined;

(vi) $\begin{pmatrix} 104 & 122 \\ 100 & 120 \\ 19 & 28 \end{pmatrix}$; (vii) $\begin{pmatrix} 5 & 11 & 17 & 23 \\ 11 & 25 & 39 & 53 \\ 17 & 39 & 61 & 83 \\ 23 & 53 & 83 & 113 \end{pmatrix}$; (viii) $\begin{pmatrix} 84 & 100 \\ 100 & 120 \end{pmatrix}$;

(ix) (919, 1012); (x) $\begin{pmatrix} 9056 & 13400 \\ 14400 & 19400 \end{pmatrix}$.

Exercise 2: (i) $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 2 & 3 \\ 3 & 4 & 1 & 2 \\ 2 & 3 & 4 & 1 \end{vmatrix} =$

$$\begin{vmatrix}1 & 2 & 3\\4 & 1 & 2\\3 & 4 & 1\end{vmatrix} - 2\begin{vmatrix}4 & 2 & 3\\3 & 1 & 2\\2 & 4 & 1\end{vmatrix} + 3\begin{vmatrix}4 & 1 & 3\\3 & 4 & 2\\2 & 3 & 1\end{vmatrix} - 4\begin{vmatrix}4 & 1 & 2\\3 & 4 & 1\\2 & 3 & 4\end{vmatrix}$$

$$= \left[\begin{vmatrix}1 & 2\\4 & 1\end{vmatrix} - 2\begin{vmatrix}4 & 2\\3 & 1\end{vmatrix} + 3\begin{vmatrix}4 & 1\\3 & 4\end{vmatrix}\right]$$

$$- 2\left[4\begin{vmatrix}1 & 2\\4 & 1\end{vmatrix} - 2\begin{vmatrix}3 & 2\\2 & 1\end{vmatrix} + 3\begin{vmatrix}3 & 1\\2 & 4\end{vmatrix}\right]$$

$$+ 3\left[4\begin{vmatrix}4 & 2\\3 & 1\end{vmatrix} - \begin{vmatrix}3 & 2\\2 & 1\end{vmatrix} + 3\begin{vmatrix}3 & 4\\2 & 3\end{vmatrix}\right]$$

$$- 4\left[4\begin{vmatrix}4 & 1\\3 & 4\end{vmatrix} - \begin{vmatrix}3 & 1\\2 & 4\end{vmatrix} + 2\begin{vmatrix}3 & 4\\2 & 3\end{vmatrix}\right]$$

$= [-7 - 2(-2) + 3(13)] - 2[4(-7) - 2(-1) + 3(10)]$
$+ 3[4(-2) - (-1) + 3(1)] - 4[4(13) - 10 + 2(1)]$

$= 36 - 2(4) + 3(-4) - 4(44)$

$= 36 - 8 - 12 + 176$

$= - 160.$

(ii) $\begin{vmatrix}1 & 2 & 3 & 4\\4 & 1 & 2 & 3\\3 & 4 & 1 & 2\\2 & 3 & 4 & 1\end{vmatrix}$

$$= \begin{vmatrix}1 & 2\\4 & 1\end{vmatrix} . \begin{vmatrix}1 & 2\\4 & 1\end{vmatrix} - \begin{vmatrix}1 & 3\\4 & 2\end{vmatrix} . \begin{vmatrix}4 & 2\\3 & 1\end{vmatrix} + \begin{vmatrix}1 & 4\\4 & 3\end{vmatrix} . \begin{vmatrix}4 & 1\\3 & 4\end{vmatrix}$$

$$+\begin{vmatrix}2 & 3\\1 & 2\end{vmatrix}\cdot\begin{vmatrix}3 & 2\\2 & 1\end{vmatrix} - \begin{vmatrix}2 & 4\\1 & 3\end{vmatrix}\cdot\begin{vmatrix}3 & 1\\2 & 4\end{vmatrix} + \begin{vmatrix}3 & 4\\2 & 3\end{vmatrix}\cdot\begin{vmatrix}3 & 4\\2 & 3\end{vmatrix}$$

$= (-7)^2 - (-10)(-2) + (-13)(13) + 1(-1) - 2(10) + 1^2$

$= 49 - 20 - 169 - 1 - 20 + 1 = -160.$

(iii) $$\begin{vmatrix}1 & 2 & 3 & 4\\4 & 1 & 2 & 3\\3 & 4 & 1 & 2\\2 & 3 & 4 & 1\end{vmatrix} = \begin{vmatrix}10 & 10 & 10 & 10\\4 & 1 & 2 & 3\\3 & 4 & 1 & 2\\2 & 3 & 4 & 1\end{vmatrix}$$

$R_1 + R_2, R_1 + R_3, R_1 + R_4$

$$= 10\begin{vmatrix}1 & 1 & 1 & 1\\4 & 1 & 2 & 3\\3 & 4 & 1 & 2\\2 & 3 & 4 & 1\end{vmatrix}$$

$R_1 \div 10$

$$= 10\begin{vmatrix}1 & 0 & 0 & 0\\4 & -3 & -2 & -1\\3 & 1 & -2 & -1\\2 & 1 & 2 & -1\end{vmatrix}$$

$C_2 - C_1, C_3 - C_1, C_4 - C_1$

$$= -10 \begin{vmatrix} 3 & 2 & 1 \\ 1 & -2 & -1 \\ 1 & 2 & -1 \end{vmatrix} \quad \text{expanding along } R_1$$

$$= -10 \begin{vmatrix} 3 & 2 & 1 \\ 1 & -2 & -1 \\ 0 & 4 & 0 \end{vmatrix} \quad R_3 - R_2$$

$$= 40 \begin{vmatrix} 3 & 1 \\ 1 & -1 \end{vmatrix} \quad \text{expanding along } R_3$$

$$= 40(-4)$$

$$= -160.$$

This determinant is an example of a cyclic determinant, whereby each row is the same as the one above, but shifted one column to the right, with the last entry moving back to the first. There's a very easy way to evaluate such determinants using eigenvalues and eigenvectors. These will be discussed in a later chapter.

Exercise 3:

(a) $\text{adj}A = \begin{pmatrix} \begin{vmatrix} 2 & 4 \\ 3 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 4 \\ 1 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} \\ -\begin{vmatrix} 1 & 1 \\ 3 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 9 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} \\ \begin{vmatrix} 1 & 1 \\ 2 & 4 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 1 & 4 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} \end{pmatrix}^T = \begin{pmatrix} 6 & -5 & 1 \\ -6 & 8 & -2 \\ 2 & -3 & 1 \end{pmatrix}^T$

$= \begin{pmatrix} 6 & -6 & 2 \\ -5 & 8 & -3 \\ 1 & -2 & 1 \end{pmatrix}$

$|A| = (1, 1, 1)\begin{pmatrix} 6 \\ -5 \\ 1 \end{pmatrix} = 2.$

So $A^{-1} = \frac{1}{2}\begin{pmatrix} 6 & -6 & 2 \\ -5 & 8 & -3 \\ 1 & -2 & 1 \end{pmatrix}.$

(b) $\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 4 & 0 & 1 & 0 \\ 1 & 3 & 9 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 2 & 8 & -1 & 0 & 1 \end{array}\right)$ $(R_2 - R_1, R_3 - R_1)$

$\to \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 2 & 1 & -2 & 1 \end{array}\right)$ $(R_3 - 2R_2)$

$\to \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & -1 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & -1 & \frac{1}{2} \end{array}\right)$ $(R_3 \div 2)$

$$\rightarrow \begin{pmatrix} 1 & 1 & 0 & 1/2 & 1 & -1/2 \\ 0 & 1 & 0 & -5/2 & 4 & -3/2 \\ 0 & 0 & 1 & 1/2 & -1 & 1/2 \end{pmatrix} \quad (R_1 - R_3,\ R_1 - 3R_3)$$

$$\rightarrow \begin{pmatrix} 1 & 0 & 0 & 3 & -3 & 1 \\ 0 & 1 & 0 & -5/2 & 4 & -3/2 \\ 0 & 0 & 1 & 1/2 & -1 & 1/2 \end{pmatrix} \quad (R_1 - R_2)$$

Hence $A^{-1} = \begin{pmatrix} 3 & -3 & 1 \\ -5/2 & 4 & -3/2 \\ 1/2 & -1 & 1/2 \end{pmatrix}$.

Exercise 4:

$$\left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 4 & 1 & 0 & 0 & 0 \\ 0 & 1 & 5 & 6 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 7 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \rightarrow \left(\begin{array}{cccc|cccc} 1 & 2 & 3 & 0 & 1 & 0 & 0 & -4 \\ 0 & 1 & 5 & 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

$(R_1 - 4R_4,\ R_2 - 6R_4,\ R_3 - 7R_4)$

$$\rightarrow \left(\begin{array}{cccc|cccc} 1 & 2 & 0 & 0 & 1 & 0 & -3 & 17 \\ 0 & 1 & 0 & 0 & 0 & 1 & -5 & 29 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \quad (R_1 - 3R_3,\ R_2 - 5R_3)$$

$$\rightarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & -2 & 7 & -41 \\ 0 & 1 & 0 & 0 & 0 & 1 & -5 & 29 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \quad (R_1 - 2R_2)$$

So the inverse is $\begin{pmatrix} 1 & -2 & 7 & -41 \\ 0 & 1 & -5 & 29 \\ 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

Exercise 5: (i) $A^T = (BB^T)^T = B^{TT}B^T = BB^T = A$, so A is symmetric.

(ii) Suppose that $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}$ is a 2×2 real matrix and $A = BB^T$ for some real matrix B.

Since $|A| = |B|.|B^T| = |B|^2$ we must have $ad - b^2 \geq 0$.

If $B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ then $\begin{pmatrix} a & b \\ b & d \end{pmatrix} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}\begin{pmatrix} e & g \\ f & h \end{pmatrix} = \begin{pmatrix} e^2 + f^2 & eg + fh \\ eg + fh & g^2 + h^2 \end{pmatrix}$

and so $a \geq 0$ and $d \geq 0$.

Now suppose that $a \geq 0$, $d \geq 0$ and $ad - b^2 \geq 0$.

Let $A = \begin{pmatrix} a & b \\ b & d \end{pmatrix} = BB^T$ where $B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ and e, f, g, h are real.

Then $A = BB^T = \begin{pmatrix} e & f \\ g & h \end{pmatrix}\begin{pmatrix} e & g \\ f & h \end{pmatrix} = \begin{pmatrix} e^2 + f^2 & eg + fh \\ eg + fh & g^2 + h^2 \end{pmatrix}$ so

$$\left.\begin{aligned} e^2 + f^2 &= a \\ eg + fh &= \mathrm{b} \\ g^2 + h^2 &= d \end{aligned}\right\}.$$

We must show that this system has real solutions.

Since we have 3 equations in 4 variables there's a chance that if we choose one of the variables to be zero we'll still get a solution.

Let $e = 0$. The equations reduce to $\left.\begin{aligned} f^2 &= a \\ fh &= b \\ g^2 + h^2 &= d \end{aligned}\right\}$.

Suppose $a = 0$. Then $f = 0$ and, since $ad - b^2 \geq 0$ we must have $b \leq 0$ and so $b = 0$.

If $a = 0$ then the fact that $ad - b^2 \geq 0$ means that $b^2 \leq 0$, and the fact that b is real means that $b = 0$.

So in this case A = $\begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix}$. Since $d \geq 0$, B = $\begin{pmatrix} 0 & 0 \\ 0 & \sqrt{d} \end{pmatrix}$ is a real matrix with $BB^T = A$.

Suppose now that $a > 0$.

Here we have $f = \sqrt{a}$ and $h = \dfrac{b}{\sqrt{a}}$. It remains for us to substitute into the equation $g^2 + h^2 = d$ and solve for g.

This gives $g^2 + \dfrac{b^2}{a} = d$ and so $g^2 = \dfrac{ad - b^2}{a}$.

Since $ad - b^2 \geq 0$ and $a > 0$ we may take $g = \sqrt{\dfrac{ad - b^2}{a}}$.

Let B = $\begin{pmatrix} 0 & \sqrt{a} \\ \sqrt{\dfrac{ad - b^2}{a}} & \dfrac{b}{\sqrt{a}} \end{pmatrix}$. Then $BB^T = A$.

Exercise 6:

$$\left(\begin{array}{ccccc|cccc} 2 & -1 & 2 & -1 & 3 & 1 & 0 & 0 & 0 \\ 4 & 1 & 6 & -5 & 8 & 0 & 1 & 0 & 0 \\ 2 & -10 & -4 & 8 & -5 & 0 & 0 & 1 & 0 \\ 2 & -13 & -6 & 16 & -5 & 0 & 0 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccccc|cccc} 2 & -1 & 2 & -1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 2 & -3 & 2 & -2 & 1 & 0 & 0 \\ 0 & -9 & -6 & 9 & -8 & -1 & 0 & 1 & 0 \\ 0 & -12 & -8 & 17 & -8 & -1 & 0 & 0 & 1 \end{array}\right)$$

$R_2 - 2R_1$, $R_3 - R_1$, $R_4 - R_1$

$$\to \left(\begin{array}{ccccc|cccc} 2 & -1 & 2 & -1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 2 & -3 & 2 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & -7 & 3 & 1 & 0 \\ 0 & 0 & 0 & 5 & 0 & -9 & 4 & 0 & 1 \end{array}\right)$$

$R_3 + 3R_2$, $R_4 + 4R_2$

$$\to \left(\begin{array}{ccccc|cccc} 2 & -1 & 2 & -1 & 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 2 & -3 & 2 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 & -9 & 4 & 0 & 1 \\ 0 & 0 & 0 & 0 & -2 & -7 & 3 & 1 & 0 \end{array}\right)$$

$R_4 + 2R_3$

So U = $\begin{pmatrix} 2 & -1 & 2 & -1 & 3 \\ 0 & 3 & 2 & -3 & 2 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}$ and E = $\begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -9 & 4 & 0 & 1 \\ -7 & 3 & 1 & 0 \end{pmatrix}$

$$E\mathbf{b} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -9 & 4 & 0 & 1 \\ -7 & 3 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3 \\ 2 \end{pmatrix}.$$

$$(U \mid E\mathbf{b}) = \left(\begin{array}{ccccc|c} 2 & -1 & 2 & -1 & 3 & 1 \\ 0 & 3 & 2 & -3 & 2 & 0 \\ 0 & 0 & 0 & 5 & 0 & 3 \\ 0 & 0 & 0 & 0 & -2 & 2 \end{array}\right)$$

By back substitution, $x_5 = -1$, $x_4 = \frac{3}{5}$, $x_3 = k$,

$3x_2 = -2k + \frac{9}{5} + 2 = \frac{19 - 10k}{5}$ so $x_2 = \frac{19 - 10k}{15}$,

$2x_1 = 1 + \frac{19 - 10k}{15} - 2k + \frac{3}{5} + 3 = \frac{88 - 40k}{15}$ so

$x_1 = \frac{88 - 40k}{30}$.