MA 106: Linear Algebra

Instructors: Prof. G. K. Srinivasan and S. Krishnan

Text and references

Main Text:
E. Kreyszig, Advanced Engineering Mathematics, 8^{th} ed. (Chapters 6 and 7)

Additional references:

1) S. Kumaresan, Linear Algebra- A geometric approach.
2) Wylie and Barrett, Advanced Engineering Mathematics, 6th ed. (Chapter 13)

More course details

Syllabus, Grading policy, etc: see the “Basic Information” file on moodle.

Study hours:

1. Lectures - 21 hours
2. Tutorials - 7 hours
3. Independent study - 28 hours

Outline of Week-1

1. Matrices
2. Addition, multiplication, transposition
3. Linear transformations and matrices
4. Linear equations and Gauss' elimination
5. Row echelon forms and elementary row matrices
6. Reduced REF
7. Gauss-Jordan method for finding inverse

Matrices

Definition 1

A rectangular array of numbers, *real or complex,* is called a matrix.

An array could be of any type of non mathematical objects too. Or more sophisticated mathematical objects like functions instead of numbers. E.g.

$$\begin{bmatrix} \sin x & -\cos x \\ \cos x & \sin x \end{bmatrix}.$$

Most of the topics today (sec 6.1 and 6.2) will be briefly reviewed and the details will be left for *self study*

Some matrices already seen

In MA-109 and MA-111 you would have seen Jacobians, derivatives, Hessians, Wronskian, ...etc.

1. If $f : \mathbb{R}^3 \mapsto \mathbb{R}$ is a differentiable function, then we have the 1×3 matrix $\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)$.
2. If $f : \mathbb{R}^3 \mapsto \mathbb{R}^2$ is a function with $f(x, y, z) = (u(x, y, z), v(x, y, z))$, then we have $Df = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{pmatrix}$. Df is clearly a 2×3 matrix.
3. If $u = r\cos\theta$ and $v = r\sin\theta$, then we have seen $J = \begin{pmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial \theta} \end{pmatrix}$.

Basic Notation

$$\boxed{A = [a_{jk}],\ 1 \leq j \leq m,\ 1 \leq k \leq n}$$

denotes an $m \times n$ matrix whose entry in j^{th} row and the k^{th} column is the number a_{jk}. (Equivalently, k^{th} entry in the j^{th} row or equivalently $j^{th} \cdots$.)

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \cdots & \boxed{a_{jk}} & \cdots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mk} & \cdots & a_{mn} \end{bmatrix}$$

Some special cases are sometimes named differently

Definition 2

An $m \times 1$ matrix is referred to as a $\boxed{\text{column vector}}$ while a $1 \times n$ matrix is referred to as a $\boxed{\text{row vector}}$.

Examples :

1. $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ is a column.
2. $[0 \;\; 1 \;\; -1 \;\; 3 \;\; 0]$ is a row.

Two matrices are said to be equal if and only if their <u>corresponding</u> entries are same. [0.5]

Transposition

Definition 3

A matrix B is a transpose of A if the rows of B are the columns of A and vice versa.

Thus, is $A = [a_{jk}]$ is an $m \times n$ then B is $n \times m$ matrix $[b_{rs}]$ where $b_{rs} = a_{sr}$; $1 \leq r \leq n$, $1 \leq s \leq m$.

The transpose of A is unique and is denoted by A^T.

Example: $A = \begin{bmatrix} 5 & -8 & 1 \\ 4 & 0 & 0 \end{bmatrix} \Longrightarrow A^T = \begin{bmatrix} 5 & 4 \\ -8 & 0 \\ 1 & 0 \end{bmatrix}$.

Exercise: Show that $(A^T)^T = A$.

Symmetry, Addition, Scalar multiplication

Definition 4

A matrix A is called symmetric (resp. skew-symmetric) if $A = A^T$ (resp. $A = -A^T$).

These are necessarily square matrices i.e.

no. of rows=no. of columns.

Let A, B be *real (or complex)* matrices and $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) be a scalar.

Definition 5

(Addition) If $A = [a_{jk}]$ and $B = [b_{jk}]$ have the same order $m \times n$, we define their addition to be $A + B = [c_{jk}] = [a_{jk} + b_{jk}]$.

Definition 6

(Scalar multiplication) The scalar multiplication of λ with A is defined as $\lambda A = [\lambda a_{jk}]$.

Matrix multiplication

More generally, when $A = [a_{jk}]$ is $m \times n$ and $B = [b_{k\ell}]$ is $n \times p$, then the product $C := A \times B$ is a well defined $m \times p$ matrix cooked by the following recipe (called *row by column multiplication*):

$$\boxed{C = [c_{j\ell}] \text{ where } c_{j\ell} = \sum_{k=1}^{n} a_{jk} b_{k\ell}; 1 \leq j \leq m, 1 \leq \ell \leq p.}$$

$$\begin{bmatrix} \\ a_{j1} & \cdots & a_{jk} & \cdots & a_{jn} \\ \\ \end{bmatrix} \begin{bmatrix} & b_{1\ell} & \\ & \vdots & \\ & b_{k\ell} & \\ & \vdots & \\ & b_{n\ell} & \end{bmatrix}$$

Associativity

Theorem 7

If A, B, C are real (or complex) matrices such that A is $m \times n$, B is $n \times p$ and C is $p \times q$, then the products AB and BC are defined and in turn the products $A(BC)$ and $(AB)C$ are also defined and the latter two are equal.

In other words:

$$\boxed{A(BC) = (AB)C}$$

Proof:

Exercise.

Transpose of a product

Theorem 8

Let A be $m \times n$ and B be $n \times p$, then AB and $B^T A^T$ are well defined and in fact

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

Proof:
Omitted.

Exercise: Let $A = \begin{bmatrix} 4 & 9 \\ 0 & 2 \\ 1 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 3 & 7 \\ 2 & 8 \end{bmatrix}$. Compute A^T, B^T, AB, $(AB)^T$, $B^T A^T$ and $A^T B^T$ to verify the claim.

Example: Dot product as a matrix product

Definition 9

Let $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$ be column vectors of the same size n. Their dot product (or inner product or scalar product) is defined as

$$\mathbf{v} \cdot \mathbf{w} = \sum_{j=1}^{n} v_j w_j.$$

It is interesting to observe that $\mathbf{v} \cdot \mathbf{w} = \mathbf{v}^T \mathbf{w}$ as a 1×1 matrix, which being symmetric also equals $\mathbf{w}^T \mathbf{v} = \mathbf{w} \cdot \mathbf{v}$.
Question: What about $\mathbf{v}\mathbf{w}^T$? Is it defined? Is it also the dot product?

Linear transformations and associated matrices

From now on the elements of $\mathbb{R}^n$ will be written as the column vectors of length n.

Definition 10

A map $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is called linear if it is of the form $f(\mathbf{x}) = a_1x_1 + a_2x_2 \cdots + a_nx_n$ for suitable constants (scalars) $a_1, a_2, ..., a_n$. (Here x_j are the entries of $\mathbf{x}$.)

If we view $A = [a_1\ a_2\ \cdots a_n]$ as a row vector, then $f(\mathbf{x}) = A\mathbf{x}$ in terms of matrix multiplication. More generally, an $m \times n$ matrix A can be viewed as a *linear* map $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ via $\mathbf{x} \mapsto A\mathbf{x}$.

This viewpoint allows us to study matrices geometrically.

[1.0]

Linear transformations and associated matrices

Show that the range of $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ as a linear map $\mathbb{R} \longrightarrow \mathbb{R}^2$ is a line through $\mathbf{0}$.

$A(t) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} [t] = \begin{bmatrix} t \\ -t \end{bmatrix} = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$, say. Thus $x(t) = t,\ y(t) = -t$ which are parametric equations of the line $x + y = 0$ through $\mathbf{0}$.

Recall that we think of vectors in $\mathbb{R}^n$ as column vectors. If A is an $m \times n$ matrix, we get $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ by $f(x) = Ax$.

If $A = \begin{pmatrix} \cos\pi/3 & -\sin\pi/3 \\ \sin\pi/3 & \cos\pi/3 \end{pmatrix}$, this gives us a function $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ given by $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\pi/3 - y\sin\pi/3 \\ x\sin\pi/3 + y\cos\pi/3 \end{pmatrix}$. This geometrically corresponds to rotating the plane $\mathbb{R}^2$ by an angle $\pi/3$ about the origin (which will be fixed).

Linear transformations

1. Let $A = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$, this gives us a function $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ given by $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+3y \\ y \end{pmatrix}$. This geometrically corresponds to a shearing transformation.

2. Let $A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$, this gives us a function $f : \mathbb{R}^2 \mapsto \mathbb{R}^2$ given by $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}$. This geometrically corresponds to a scaling transformation. We can think of this as enlarging/shrinking a picture without messing up the aspect ratio (depending on whether $\lambda > 1$ or $\lambda < 1$ respectively).

Linear transformations and associated matrices

Determine domain and range and also show that $\begin{bmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{bmatrix}$ has a plane through $\mathbf{0}$ as its range.

(i) The matrix is 3×2, so that's a linear transformation $\mathbb{R}^2 \longrightarrow \mathbb{R}^3$.

(ii) $B\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u - v \\ -u + 2v \\ v \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, say. Thus
$x = u - v,\ y = 2v - u,\ z = v$ are the parametric equations of a plane through $\mathbf{0}$ whose equation, on eliminating u, v, is $x + y - z = 0$.

Consider the matrix: $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Determine the images of
(i) Unit square $\{0 \leq x \leq 1,\ 0 \leq y \leq 1\}$,
(ii) Unit circle $\{x^2 + y^2 = 1\}$ and
(iii) Unit disc $\{x^2 + y^2 \leq 1\}$ under the above matrices viewed as linear maps $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$.

(i) Unit square $\{0 \leq x \leq 1,\ 0 \leq y \leq 1\}$ has vertices $\{(0,0)^T, (1,0)^T, (1,1)^T, (0,1)^T\}$. Therefore the image has vertices $\{A(0,0)^T, A(1,0)^T, A(1,1)^T, A(0,1)^T\}$. The full image is a parallelogram $\{(0,0)^T, (1,0)^T, (2,1)^T, (1,1)^T\}$.

(ii) $A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix} \implies \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u-v \\ v \end{bmatrix}$ and $x^2 + y^2 = 1 \implies u^2 - 2uv + 2v^2 = 1$ which is an ellipse.

(iii) Elliptic disc enclosed by (ii).

Linear transformations and associated matrices

Consider the matrix: $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ Determine the images of
(i) Unit square $\{0 \leq x \leq 1,\ 0 \leq y \leq 1\}$,
(ii) Unit circle $\{x^2 + y^2 = 1\}$ and
(iii) Unit disc $\{x^2 + y^2 \leq 1\}$ under the above matrices viewed as linear maps $\mathbb{R}^2 \longrightarrow \mathbb{R}^2$.

(i) Unit square $\{0 \leq x \leq 1,\ 0 \leq y \leq 1\}$ has vertices $\{(0,0)^T, (1,0)^T, (1,1)^T, (0,1)^T\}$. Therefore the image has 'vertices' $\{(0,0)^T, (1,1)^T, (2,2)^T, (1,1)^T\}$ which are collinear. The full image is a line segment $(0,0)^T$ to $(2,2)^T$.

(ii) $A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ x+y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}$ and
$x^2 + y^2 = 1 \Longrightarrow \min(x+y) = -\sqrt{2},\ \max(x+y) = \sqrt{2} \Longrightarrow$ a line segment from $-\begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix}$ to $\begin{bmatrix} \sqrt{2} \\ \sqrt{2} \end{bmatrix}$.

(iii) Elliptic disc enclosed by (ii). [1.5]

Linear systems and matrices

Consider m linear equations in n variables $x_1, x_2, ..., x_n$:

$$\begin{array}{rcll} a_{11}x_1 + \; a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 & \qquad (1) \\ a_{21}x_1 + \; a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 & \qquad (2) \\ \vdots \qquad & \vdots & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m & \qquad (m) \end{array}$$

In matrix language, we can write

$$A\mathbf{x} = \mathbf{b}$$

1. where $A = [a_{jk}]$ is the $m \times n$ *coefficient matrix*,
2. $\mathbf{x} \in \mathbb{R}^n$ is the *unknown vector* and $\mathbf{b} \in \mathbb{R}^m$ is a given i.e. a *known vector*.
3. Also $\boxed{A^+ = [A|\mathbf{b}]}$ is an $m \times (n+1)$ matrix, called the *augmented* matrix and it completely describes the above system of equations.

Gauss' elimination

In order to solve the above system of linear equations, Gauss proposed 3 kinds of operations each of which gives a new linear system, (hopefully) simpler and *equivalent* to the original one.

1. Interchanging two different equations, say equations (i) and (j).
2. Multiplying an equation by a scalar and adding it to some other equation of the system.
3. Multiplying an equation by a *non-zero* number.

Each of the above operations is equivalent to an obvious corresponding operation on the rows of the augmented matrix A^+.

Elementary row operations

Let A be any matrix. The elementary row operations on A are the following:

1. Applying P_{jk} to A. This means Interchanging the j^{th} and the k^{th} rows.
2. Applying $E_{jk}(c)$ to A for $j \neq k$. This means multiplying the k^{th} row by c and adding it to the j^{th} row.
3. And lastly, applying $M_j(c)$, $c \neq 0$ to A. This implies that we multiply the j^{th} row by a scalar $c \neq 0$.

Exercise: Show that for $j \neq k$, $P_{jk} = E_{jk}(1)E_{kj}(-1)M_j(-1)E_{jk}(-1)$ i.e. we apply successively $E_{jk}(-1)$, $M_j(-1)$, $E_{kj}(-1)$ and $E_{jk}(1)$ (note the right to left order of multiplication) to obtain the effect of applying P_{jk}.

• Question: Does it mean that P_{jk} is not elementary? Obviously, we will prefer to apply P_{jk} directly.

Row echelon form

Any matrix after a sequence of ERO's (done cleverly) gets reduced to what is known as a *Row Echelon form*. In this form each row, except perhaps the first, starts with a string of zeroes. Each row

starts with strictly more number of zeroes than the previous row.

Note: The first non-zero entry in the j^{th} row is known as the j^{th} *pivot*.
The j^{th} pivot is below and strictly to the right of $(j-1)^{th}$ pivot. All the entries *below* a pivot are zeroes.
Conclusion: The no. of pivots in a *REF* of A

$\leq$ the no. of rows in A.

Example: The matrix $\begin{bmatrix} \boxed{3} & 2 & 1 & 3 \\ 0 & \boxed{1} & -1 & 6 \\ 0 & 0 & 0 & \boxed{12} \\ 0 & 0 & 0 & 0 \end{bmatrix}$ is in REF. The three pivots are indicated.

Row echelon form ctd.

The matrix $\begin{bmatrix} \boxed{3} & 2 & 1 & 3 \\ 0 & \boxed{1} & -1 & 6 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{12} \end{bmatrix}$ is NOT in REF.

Properties of the REF:

- Given a matrix A, its REF is NOT unique. However, the position of each of its pivots is unalterable.
- **Reduced REF:** A matrix in REF can be further row-operated upon to ensure that (i) each pivot becomes 1 and (ii) all the entries above each pivot become 0. This is the *reduced* REF and it is unique.
- Reduced REF is mainly of theoretical interest only.

Example 1

Reduce the augmented matrix to the REF:

$$\begin{aligned} 3x_1 + 2x_2 + x_3 &= 3 \\ 2x_1 + x_2 + x_3 &= 0 \\ 6x_1 + 2x_2 + 4x_3 &= 6 \end{aligned}$$

$$A^+ = \left[\begin{array}{ccc|c} 3 & 2 & 1 & 3 \\ 2 & 1 & 1 & 0 \\ 6 & 2 & 4 & 6 \end{array}\right] \overset{Row\ Ops}{\longmapsto} \left[\begin{array}{ccc|c} \boxed{3} & 2 & 1 & 3 \\ 0 & \boxed{1} & -1 & 6 \\ 0 & 0 & 0 & \boxed{12} \end{array}\right]$$

The matrix on the right is a REF of A^+. The pivots are highlighted. The third equation of the equivalent system of equations reads:

$$0x_1 + 0x_2 + 0x_3 = 12$$

which is absurd. The given system is inconsistent and hence has no solutions.

Example 2

Solve the system of linear equations:

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ -x_1 + x_2 - x_3 &= 0 \\ 10x_2 + 25x_3 &= 90 \\ 20x_1 + 10x_2 &= 80. \end{aligned}$$

$$A^+ = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 10 & 25 & 90 \\ 20 & 10 & 0 & 80 \end{bmatrix} \overset{\textit{Row Ops}}{\longmapsto} \left[\begin{array}{ccc|c} \boxed{-1} & -1 & 1 & 0 \\ 0 & \boxed{10} & 25 & 90 \\ 0 & 0 & \boxed{-95} & -190 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Example 2, details

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 10 & 25 & 90 \\ 20 & 10 & 0 & 80 \end{array}\right] \xmapsto{R_2(+R_1),R_4(-20R_1)} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 30 & -20 & 80 \end{array}\right]$$

$$\xmapsto{R_2 \sim R_4} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 30 & -20 & 80 \\ 0 & 10 & 25 & 90 \\ 0 & 0 & 0 & 0 \end{array}\right] \xmapsto{-R_2(+3R_3)} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 0 & 95 & 190 \\ 0 & 10 & 25 & 90 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xmapsto{R_2 \sim R_3} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 0 & 95 & 190 \\ 0 & 0 & 0 & 0 \end{array}\right] \xmapsto{R_2/5,R_3/95} \left[\begin{array}{ccc|c} \boxed{1} & -1 & 1 & 0 \\ 0 & \boxed{2} & 5 & 18 \\ 0 & 0 & \boxed{1} & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Example 2 contd.

Equations become

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ 2x_2 + 5x_3 &= 18 \\ x_3 &= 2 \\ 0 = 0. & \end{aligned}$$

The last equation shows *consistency* and the solution(s) may be obtained by the *back substitution* method as

$$x_3 = 2 \Longrightarrow x_2 = 4 \Longrightarrow x_1 = 2.$$

The solution is unique.

Example 3

Solve the system of linear equations:

$$\begin{aligned} 3x_1 + 2x_2 + 2x_3 - 5x_4 &= 8 \\ 0.6x_1 + 1.5x_2 + 1.5x_3 - 5.4x_4 &= 2.7 \\ 1.2x_1 - 0.3x_2 - 0.3x_3 + 2.4x_4 &= 2.1. \end{aligned}$$

$$A^+ = \left[\begin{array}{cccc|c} 3 & 2 & 2 & -5 & 8 \\ 0.6 & 1.5 & 1.5 & -5.4 & 2.7 \\ 1.2 & -0.3 & -0.3 & 2.4 & 2.1 \end{array}\right]$$

$$\xrightarrow{Row\ Ops} \left[\begin{array}{cccc|c} \boxed{3} & 2 & 2 & -5 & 8 \\ 0 & \boxed{1.1} & 1.1 & -4.4 & 1.1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Example 3 contd.

The equivalent system is:

$$
\begin{aligned}
3.0x_1 + 2.0x_2 + 2.0x_3 - 5.0x_4 &= 8.0 \\
1.1x_2 + 1.1x_3 - 4.4x_4 &= 1.1 \\
0 &= 0 \quad \text{(consistency)}.
\end{aligned}
$$

Solution set, by the back substitution method, is

$$
\begin{aligned}
x_2 &= 1 - x_3 + 4x_4 \\
x_1 &= 2 - x_4; \quad (x_3, x_4) \in \mathbb{R}^2.
\end{aligned}
$$

The solution set is a 2-parameter family! [2.0]

Elementary row matrices

Recall the elementary row operations on matrices:

1. P_{jk}: Interchanging the j^{th} and the k^{th} rows.
2. $E_{jk}(c)$: Adding c times the k^{th} row to the j^{th} row.
3. And $M_j(\lambda)$: Multiplying the j^{th} row by a scalar $\lambda \neq 0$.

Denote by $\mathcal{E}_{jk}$ the *standard basic matrix* whose $(j,k)^{th}$ entry is 1 and the rest 0's. We need only $m \times m$ or square matrices in the following discussion. Let $\mathbf{I} = \sum_{j=1}^{m} \mathcal{E}_{jj}$ be the $m \times m$ identity matrix.

Elementary row matrices, contd.

For a scalar c and $1 \leq j \neq k \leq m$ let us define

1. The matrix $P_{jk} = \sum_{\ell \neq j,k} \mathcal{E}_{\ell\ell} + \mathcal{E}_{jk} + \mathcal{E}_{kj}$ i.e. the matrix obtained by interchanging the j^{th} and the k^{th} rows of $\mathbf{I}$.
2. The matrix $E_{jk}(c) = \mathbf{I} + c\mathcal{E}_{jk},\ j \neq k$ i.e. the matrix obtained from $\mathbf{I}$ by adding c times the k^{th} row to the j^{th} row.
3. The matrix $M_j(\lambda) = \mathcal{E}_{11} + \mathcal{E}_{22} + \cdots \lambda\mathcal{E}_{jj} + \cdots + \mathcal{E}_{mm}$ i.e. the matrix obtained from $\mathbf{I}$ by multiplying its j^{th} row by $\lambda(\neq 0)$ in this case.

The above matrices are known as the $\boxed{\text{elementary row matrices}}$ (of order m if you like.)

Examples of $ERM's$

Among 3×3 matrices we notice that

$$\begin{aligned}
P_{12} &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1. \end{bmatrix} \\
E_{32}(c) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & c & 1. \end{bmatrix} \\
M_1(\lambda) &= \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1. \end{bmatrix}
\end{aligned}$$

$ERM's$ contd.

Theorem 11

Let A be any $m \times n$ matrix and P_{jk}, $E_{jk}(c)$ and $M_j(\lambda)$ be the $m \times m$ ERM's ($j \neq k, \lambda \neq 0$).

1. *The product $P_{jk}A$ is the $m \times n$ matrix obtained by interchanging the j^{th} and the k^{th} rows of A.*
2. *The product $E_{jk}(c)A$ is the $m \times n$ matrix obtained by adding c times the k^{th} row of A to the j^{th} row of A.*
3. *The product $M_j(\lambda)A$ is the $m \times n$ matrix obtained by multiplying the j^{th} row of A by λ.*

Proof:

Direct check. $\mathcal{E}_{jk}A$ is the matrix obtained from A by killing all its rows except the k^{th} and then moving it to the j^{th} row.

Reduced Row Echelon form

Theorem 12

Let A be an $m \times n$ matrix. There exist ERM's $E_1, E_2, ..., E_N$ of order m such that the product $E_N \cdots E_2 E_1 A$ is a row echelon form of A.

Proof:

Obvious.

Reduced Row Echelon form: Once an echelon form of A is obtained, we can by further row operations ensure that (i) each pivot becomes 1 and (ii) all the entries above each pivot vanish. This is called the *Reduced Row Echelon form* of A and is unique. [2.5]

Reduced Row Echelon form contd.

Theorem 13

Let A be a square matrix, say $n \times n$. There exist ERM's $E_1, E_2, ..., E_N$ of order n such that the product $E_N \cdots E_2 E_1 A$ is either the $n \times n$ identity matrix $\mathbf{I}$ or its last row is 0.

Proof:

Consider the *reduced* row echelon form of A. Recall that there must be $p \leq n$ pivots in all. If there are $p = n$ pivots then the *reduced* REF must be $\mathbf{I}$. If there are $p < n$ pivots,then the last $n - p$ rows must vanish.

Inverse of a square matrix

Definition 14

(Inverse of a square matrix) Let A be a square matrix. Its inverse is another square matrix B, if it exist, satisfying $AB = BA = \mathbf{I}$. We say that A is *invertible.*

Properties:

1. If A has an inverse, then it is unique. It is denoted by A^{-1}.
2. If A and B are invertible, then so is AB and $(AB)^{-1} = B^{-1}A^{-1}$.
3. Each ERM is invertible. In fact (i) $P_{jk}^{-1} = P_{jk}$, (ii) $E_{jk}(c)^{-1} = E_{jk}(-c)$ and (iii) $M_j(\lambda)^{-1} = M_j(1/\lambda)$ show that the inverses are also ERM's

The converse of the above properties is also true.

Theorem 15

A square matrix A is invertible if and only if it is a product of ERM's.

Proof:

The reduced REF $E_N \cdots E_2 E_1 A$ is either $\mathbf{I}$ or the last row is 0. In the first case $A = E_1^{-1} E_2^{-1} \cdots E_N^{-1}$ is a product of ERM's and hence itself invertible, while in the latter case
$AB = \mathbf{I} \Longrightarrow E_N \cdots E_2 E_1 AB = E_N \cdots E_2 E_1$ with the LHS having the last row vanishing.
Next, applying further row operations (without involving the last row) on the left we get the reduced REF with the last row vanishing. OTOH, the reduced REF of the RHS is clearly $\mathbf{I}$ and by the uniqueness of the reduced REF, we arrive at a contradiction.

Gauss Jordan Method

Example: Write $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ as a product of 4 elementary row matrices if $D = ad - bc \neq 0$.

Case 1 ($a \neq 0$): $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a & b \\ 0 & D/a \end{bmatrix} \mapsto \begin{bmatrix} 1 & b/a \\ 0 & 1 \end{bmatrix} \mapsto \mathbf{I}_2$. Hence

$$\begin{aligned} A &= \begin{bmatrix} 1 & 0 \\ c/a & 1 \end{bmatrix} \begin{bmatrix} 1/a & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & D/a \end{bmatrix} \begin{bmatrix} 1 & b/a \\ 0 & 1 \end{bmatrix} \\ &= E_{21}(c/a) M_1(a) M_2(D/a) E_{12}(b/a). \end{aligned}$$

Case 2: $(a = 0) \implies -D = bc \neq 0$: Then

$$A = P_{12} M_2(b) M_1(c) E_{12}(d/c).$$

Example: Find the inverse of $A = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}$.

Solution:

$$\begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{E_{21}(3), E_{31}(-1)} \begin{bmatrix} -1 & 1 & 2 \\ 0 & 2 & 7 \\ 0 & 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{E_{32}(-1)} \begin{bmatrix} -1 & 1 & 2 \\ 0 & 2 & 7 \\ 0 & 0 & -5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -4 & -1 & 1 \end{bmatrix}$$

Example contd.

$$\xrightarrow{M_1(-1),M_2(1/2),M_3(-1/5)} \begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & 7/2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 3/2 & 1/2 & 0 \\ 4/5 & 1/5 & -1/5 \end{bmatrix}$$

$$\xrightarrow{E_{13}(2),E_{23}(-7/2)} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3/5 & 2/5 & -2/5 \\ -13/10 & -1/5 & 7/10 \\ 4/5 & 1/5 & -1/5 \end{bmatrix}$$

$$\xrightarrow{E_{12}(1)} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -7/10 & 1/5 & 3/10 \\ -13/10 & -1/5 & 7/10 \\ 4/5 & 1/5 & -1/5 \end{bmatrix}$$

Example contd.

It follows that

$$A^{-1} = \frac{1}{10}\begin{bmatrix} -7 & 2 & 3 \\ -13 & -2 & 7 \\ 8 & 2 & -2 \end{bmatrix}.$$

Also putting all the row ops together,

$$\begin{aligned} A^{-1} = E_{12}(1)E_{13}(2)E_{23}(-7/2)M_1(-1) \\ \times M_2(1/2)M_3(-1/5)E_{32}(-1)E_{21}(3)E_{31}(-1) \end{aligned}$$

as a product of ERM's. [3.0]