A System of Linear Equations and Matrices

A.1 Matrices

Definition A.1 A *matrix* is an $m \times n$ rectangular array of numbers. The numbers in the array are called the *entries* in the matrix.

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

It is called an $m \times n$ matrix, a matrix with m rows and n columns, it is also denoted by $A = [a_{ij}]$. Two matrices are defined to be *equal* if they have the same size and their corresponding entries are equal. The entry $a_{i,j}$ in the i-th row j-th column of a matrix A is denoted by $(A)_{i,j}$.

An $n \times n$ matrix is called a *square matrix*. A square matrix with 1's on the main diagonal and 0's off the main diagonal is called an *identity matrix* and is denoted by I, or I_n when it is of size $n \times n$.

Definition A.2 Let A and B be matrices of the same size and c a scalar. Then the *sum* $A + B$ is the matrix obtained by adding the entries of B to the corresponding entries of A. The the *product* cA is the matrix obtained by multiplying each entry of the matrix A by c. The matrix cA is said to be a it scalar multiple of A.

$$A+B = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ & \cdots\cdots\cdots & & \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix}, cA = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ & \cdots\cdots\cdots & & \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$

Definition A.3 If $A = (a_{i,j})$ is an $m \times r$ matrix and $B = (b_{k,l})$ is an $r \times n$ matrix, then the *product* $C = AB$ is the $m \times n$ matrix whose (s,t) entry $c_{s,t}$ is defined as follows.

$$c_{s,t} = \sum_{u=1}^{r} a_{s,u}b_{u,t} = a_{s,1}b_{1,t} + a_{s,2}b_{2,t} + \cdots + a_{s,r}b_{r,t}.$$

$$C = AB = \begin{bmatrix} \sum_{u=1}^{r} a_{1,u}b_{u,1} & \sum_{u=1}^{r} a_{1,u}b_{u,2} & \cdots & \sum_{u=1}^{r} a_{1,u}b_{u,n} \\ \sum_{u=1}^{r} a_{2,u}b_{u,1} & \sum_{u=1}^{r} a_{2,u}b_{u,2} & \cdots & \sum_{u=1}^{r} a_{2,u}b_{u,n} \\ & \cdots\cdots\cdots & & \\ \sum_{u=1}^{r} a_{m,u}b_{u,1} & \sum_{u=1}^{r} a_{m,u}b_{u,2} & \cdots & \sum_{u=1}^{r} a_{m,u}b_{u,n} \end{bmatrix}$$

Proposition A.1 *Let A be an $m \times r$ matrix and $B = [\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n]$ be an $r \times n$ matrix whose j-th column is $\boldsymbol{b}_j$. Then*

$$AB = A[\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n] = [A\boldsymbol{b}_1, A\boldsymbol{b}_2, \ldots, A\boldsymbol{b}_n].$$

Definition A.4 If A is an $m \times n$ matrix, then the *transpose* of A, denoted by A^T, is defined to be the $n \times m$ matrix that results from interchanging the rows and columns of A, that is $(A^T)_{i,j} = A_{j,i}$ $(1 \leq i \leq n,\ 1 \leq j \leq m)$.

Definition A.5 If A is a square matrix, and if a matrix B of the same size can be found such that $AB = BA = I$, then A is said to be *invertible* and B is called the inverse of A. If not such matrix B can be found, then A is said to be *singular*.

Proposition A.2 (a) *If A is an $m \times r$ matrix and B is an $r \times n$ matrix, then $(AB)^T = B^T A^T$.*

(b) *If both A and B are invertible matrices. Then AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*

(c) *If A is an invertible matrix, then A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$.*

A.2 System of Linear Equations

Definition A.6 An arbitrary system of m linear equations in n unknowns can be written as

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{cases}$$

where $x_1, x_2, \ldots, x_n$ are the unknowns. If $b_1 = b_2 = \cdots = b_m = 0$, the system is called *homogeneous*. The *augmented matrix A* or extended coefficient matrix, and the *coefficient matrix C* of this system are defined as follows.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}, \; C = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots\cdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Definition A.7 An $m \times n$ matrix A is in *reduced row-echelon form* if the following conditions hold:

1. If a row does not consist entirely of zeros, then the first nonzero number in the row is a 1. We call this a *leading* 1.

2. If there are any rows that consist entirely zeros, then they are grouped together at the bottom of the matrix.

3. In any successive rows that do not consist entirely of zeros, the leading 1 in the lower row occurs farther to the right than the leading 1 in the higher row.

4. Each column that contains a leading 1 has zeros everywhere else in that column.

If the conditions 1, 2 and 3 are satisfied A is said to be in *row-echelon form*.

Definition A.8 The following are called *elementary row operations*.

1. $[i;c]$: Multiply row i through by a nonzero constant c.
2. $[i,j]$: Interchange rows i and j.
3. $[i,j;c]$: Add c times row i to row j.

Definition A.9 An $n \times n$ matrix is called an *elementary matrix* if it can be obtained from the $n \times n$ identity matrix I_n by performing a single elementary row operation.

1. $P(i;c)$: the matrix obtained from I_n by performing $[i;c]$ $(c \neq 0)$.
2. $P(i,j)$: the matrix obtained from I_n by performing $[i,j]$.
3. $P(i,j;c)$: the matrix obtained from I_n by performing $[i,j;c]$.

Proposition A.3 (1.5.1) *If the elementary matrix E results from performing a certain row operation on I_m and A is an $m \times n$ matrix, then the product EA is the matrix that results when this same row operation is performed on A.*

Proposition A.4 (1.5.2) *Every elementary matrix is invertible, and the inverse is also an elementary matrix.*

$$P(i;c)^{-1} = P(i;1/c),\ P(i,j)^{-1} = P(i,j),\ \textit{and}\ P(i,j;c)^{-1} = P(i,j;-c).$$

Proposition A.5 (1.6.4) *If A is an $n \times n$ matrix, then the following are equivalent.*

(a) *A is invertible.*

(b) *$A\boldsymbol{x} = \mathbf{0}$ has only the trivial solution, i.e., $\boldsymbol{x} = \mathbf{0}$.*

(c) *The reduced row-echelon form of A is I_n.*

(d) *A is expressible as a product of elementary matrices.*

(e) *$A\boldsymbol{x} = \boldsymbol{b}$ is consistent for every $n \times 1$ matrix $\boldsymbol{b}$.*

(f) *$A\boldsymbol{x} = \boldsymbol{b}$ has exactly one solution for every $n \times 1$ matrix $\boldsymbol{b}$.*

Proposition A.6 *Let A and B be $m \times n$ matrices. Then the following are equivalent.*

(a) *B is obtained by a successive application of elementary row operations.*

(b) *There is an invertible matrix P of size m such that $B = PA$.*

Theorem A.7 (Gauss-Jordan Elimination) *Let A be an $m \times n$ matrix. Then the following hold.*

(a) *By a successive application of elementary row operations, A can be reduced to a reduced row-echelon matrix.*

(b) *There is a sequence $E_1, E_2, \ldots, E_t$ of elementary matrices of size m such that $E_t E_{t-1} \cdots E_2 E_1 A$ is a reduced row-echelon matrix.*

(c) *There is an invertible matrix P such that PA is a reduced row-echelon matrix.*

Proof. (a) is explained in Section 1.2. (b) is obtained from (a) by Proposition A.3. Finally (c) is from Proposition A.5 as an invertible matrix is a product of elementary matrices. ∎

Theorem A.8 (1.2.1) *A homogeneous system of linear equations with more unknowns than equations has infinitely many solutions. In particular, if A is an $m \times n$ matrix, then a system of linear equation $A\boldsymbol{x} = \mathbf{0}$ has a nonzero solution.*

Theorem A.9 *Let A be an $n \times n$ square matrix and $I = I_n$ be the identity matrix of size n. Then the reduced row echelon form of a matrix $[A, I]$ is of the form $[I, B]$, and $B = A^{-1}$.*

Example A.1 The right hand side matrix is the augmented matrix of the system of linear equations on the left below.

$$\left\{\begin{array}{rcr} x_1 + 0x_2 + x_3 + 0x_4 + x_5 + 3x_6 & = & -1 \\ -x_1 + 0x_2 - x_3 + 0x_4 + 0x_5 - 4x_6 & = & -1 \\ 0x_1 + x_2 - 2x_3 + 3x_4 + 0x_5 - x_6 & = & 3 \\ -2x_1 - 2x_2 + 2x_3 - 6x_4 - 2x_5 - 4x_6 & = & -4 \end{array}\right. \qquad \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 3 & -1 \\ -1 & 0 & -1 & 0 & 0 & -4 & -1 \\ 0 & 1 & -2 & 3 & 0 & -1 & 3 \\ -2 & -2 & 2 & -6 & -2 & -4 & -4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 3 & -1 \\ -1 & 0 & -1 & 0 & 0 & -4 & -1 \\ 0 & 1 & -2 & 3 & 0 & -1 & 3 \\ -2 & -2 & 2 & -6 & -2 & -4 & -4 \end{bmatrix} \xrightarrow{[4,1;2]} \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 3 & -1 \\ -1 & 0 & -1 & 0 & 0 & -4 & -1 \\ 0 & 1 & -2 & 3 & 0 & -1 & 3 \\ 0 & -2 & 4 & -6 & 0 & 2 & -6 \end{bmatrix} \xrightarrow{[2,1;1]}$$

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 3 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & 1 & -2 & 3 & 0 & -1 & 3 \\ 0 & -2 & 4 & -6 & 0 & 2 & -6 \end{bmatrix} \xrightarrow{[2,3]} \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 3 & -1 \\ 0 & 1 & -2 & 3 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & -2 & 4 & -6 & 0 & 2 & -6 \end{bmatrix} \xrightarrow{[4,2;2]}$$

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 3 & -1 \\ 0 & 1 & -2 & 3 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{[1,3;-1]} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 4 & 1 \\ 0 & 1 & -2 & 3 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left\{\begin{array}{rcc} x_1 & = & 1 - s - 4u, \\ x_2 & = & 3 + 2s - 3t + u, \\ x_3 & = & s, \\ x_4 & = & t, \\ x_5 & = & -2 + u, \\ x_6 & = & u. \end{array}\right. \quad \text{or} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ 0 \\ -2 \\ 0 \end{bmatrix} + s \cdot \begin{bmatrix} -1 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \cdot \begin{bmatrix} 0 \\ -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + u \cdot \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

s, t and u are parameters.