INTRODUCTION TO MATRIX ALGEBRA

1. Definition of a matrix and a vector

1.1. **Definition of a matrix.** A matrix is a rectangular array of numbers arranged into rows and columns. It is written as

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{1}$$

The above array is called an m by n (m $\times$ n) matrix since it has m rows and n columns. Typically upper-case letters are used to denote a matrix and lower case letters with subscripts the elements. The matrix A is also often denoted

$$A = \|a_{ij}\| \tag{2}$$

Consider the following 3$\times$3 example.

$$\begin{pmatrix} 2 & 3 & -2 \\ -1 & 4 & 13 \\ 4 & \frac{2}{3} & -1 \end{pmatrix} \tag{3}$$

In this matrix a_{31} = 4 and a_{23} = 13.

1.2. **Definition of a vector.** A vector is a n-tuple of numbers. In two dimensional space or R^2, a vector would be an ordered pair of numbers $\{x, y\}$. In three dimensional space or R^3, a vector is a 3-tuple, i.e., $\{x_1, x_2, x_3\}$. Similarly for R^n. Vectors are usually denoted by lower case letters such as a or b, or more formally $\vec{a}$ or $\vec{b}$.

1.3. **Row and column vectors.**

1.3.1. *Row vector.* A matrix with one row and n columns (1$\times$n) is called a row vector. It is usually written $\vec{x}'$ or

$$\vec{x} = \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_n \end{pmatrix} \tag{4}$$

The use of the prime $'$ symbol indicates we are writing the n-tuple horizontally as if it were the row of a matrix. Note that each row of a matrix is a row vector. A row vector might be as follows

$$\vec{z} = \begin{pmatrix} -1 & 4 & 13 \end{pmatrix} \tag{5}$$

where z_2 = 4.

Date: 4 March 2008.

1.3.2. *Column vector.* A matrix with one column and n rows (n×1) is called a column vector. It is written as

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ . \\ . \\ . \\ x_n \end{pmatrix} \tag{6}$$

A column vector might be as follows

$$\vec{p} = \begin{pmatrix} -2 \\ 13 \\ -1 \end{pmatrix} \tag{7}$$

Note that each column of a matrix is a column vector. It is common to write the columns of a matrix as $a_1, a_2, \ldots a_n$ where each column vector a_j is of length m. As an example a_2 is given by

$$\vec{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \\ . \\ . \\ . \\ a_{m2} \end{pmatrix} \tag{8}$$

In equation 3, a_2 is given by

$$\vec{a}_2 = \begin{pmatrix} 3 \\ 4 \\ \frac{2}{3} \end{pmatrix} \tag{9}$$

2. Various types of Matrices and Vectors

2.1. **Square matrices.** A square matrix is a matrix with an equal number of rows and columns, i.e. m=n.

2.2. **Transpose of a matrix.** The transpose of a matrix A is a matrix formed from A by interchanging rows and columns such that row i of A becomes column i of the transposed matrix. The transpose is denoted by A' or A^T and

$$A' = \|a_{ji}\| \; when \; A = \|a_{ij}\| \tag{10}$$

If a'_{ij} is the ij_{th} element of A', then $a'_{ij} = a_{ji}$. If the matrix A is given by

$$A = \begin{pmatrix} 3 & 2 & 5 & 7 \\ 1 & 4 & 6 & 3 \\ 5 & 10 & -2 & 0 \\ 1 & 1 & 15 & -2 \end{pmatrix} \tag{11}$$

then A' is given by

$$A' = \begin{pmatrix} 3 & 1 & 5 & 1 \\ 2 & 4 & 10 & 1 \\ 5 & 6 & -2 & 15 \\ 7 & 3 & 0 & -2 \end{pmatrix} \tag{12}$$

2.3. **Symmetric matrix.** A symmetric matrix is a square matrix A for which

$$A = A' \tag{13}$$

An example of a symmetric matrix is

$$\begin{aligned} T &= \begin{pmatrix} 3 & 1 & 5 & 1 \\ 1 & 4 & 10 & 1 \\ 5 & 10 & -2 & 15 \\ 1 & 1 & 15 & -2 \end{pmatrix} \\ T' &= \begin{pmatrix} 3 & 1 & 5 & 1 \\ 1 & 4 & 10 & 1 \\ 5 & 10 & -2 & 15 \\ 1 & 1 & 15 & -2 \end{pmatrix} \end{aligned} \tag{14}$$

2.4. **Identity matrix.** The identity matrix of order n written I or I_n, is a square matrix having ones along the main diagonal (the diagonal running from upper left to lower right and zeroes elsewhere).

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix} \tag{15}$$

If we write I = $\| \delta_{ij} \|$ then

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{16}$$

The symbol δ_{ij} is called the Kronecker delta.

2.5. **Scalar matrix.** For any scalar λ, the square matrix

$$S = \| \lambda \, \delta_{ij} \| = \lambda I \tag{17}$$

is called a scalar matrix. An example is

$$\begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \tag{18}$$

2.6. **Diagonal matrix.** A square matrix

$$D = \| \lambda_i \delta_{ij} \| \tag{19}$$

is called a diagonal matrix. Notice that λ_i varies with i. An example is

$$\begin{pmatrix} 13 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 56 \end{pmatrix} \tag{20}$$

If a system of equations in four variables was written with this coefficient matrix, we could solve the system by solving each equation individually because each variable would appear in each equation only once.

2.7. **Null or zero matrix.** The null or zero matrix is a matrix with each element being zero. It is denoted as **0**.

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix} \tag{21}$$

2.8. **Upper triangular matrix.** A matrix with all elements below the main diagonal equal to zero is called an upper triangular matrix.

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ 0 & a_{22} & a_{23} & ... & a_{2n} \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & a_{mn} \end{pmatrix} \tag{22}$$

Specifically $a_{ij} = 0$ if $i > j$ as long as $i < m$ and $j < n$.

2.9. **Lower triangular matrix.** A matrix with all elements above the main diagonal equal to zero is called a lower triangular matrix.

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \dots & 0 \\ a_{21} & a_{22} & 0 & \dots & 0 \\ a_{31} & a_{32} & a_{33} & \dots & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix} \tag{23}$$

Specifically $a_{ij} = 0$ if $i < j$ as long as $i < m$ and $j < n$.

The following two matrices are upper triangular and lower triangular respectively.

$$\begin{pmatrix} 3 & -1 & 7 & 6 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 2 & \frac{2}{3} \\ 0 & 0 & 0 & 6 \end{pmatrix} \qquad \begin{pmatrix} 3 & 0 & 0 & 0 \\ 4 & -1 & 0 & 0 \\ -3 & \frac{1}{3} & -4 & 0 \\ 5 & 2 & -1 & -2 \end{pmatrix} \tag{24}$$

3. A NOTE ON SUMMATION NOTATION

3.1. **Single sums.**

3.1.1. *Definition of a single sum.*

$$\sum_{i=m}^{n} a_i = a_m + a_{m+1} + a_{m+2} + ... + a_n \tag{25}$$

For example, suppose we have a vector with the following elements

$$\begin{aligned} \vec{a} &= \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & \dots & a_9 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 3 & -2 & 6 & 2 & -1 & 4 & 3 & 5 \end{pmatrix} \end{aligned} \tag{26}$$

Then

$$\sum_{i=3}^{6} a_i = -2 + 6 + 2 + -1 = 5 \tag{27}$$

3.1.2. *Properties of a single sum.*

$$\begin{aligned} \sum_{i=1}^{n} k a_i &= k \sum_{i=1}^{n} a_i \\ \sum_{i=1}^{n} k &= k + k + k + ... + k = nk \\ \sum_{i=1}^{n} (a_i + b_i) &= \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i \end{aligned} \tag{28}$$

3.2. **Double sums.**

3.2.1. *Definition of a double sum.*

$$\begin{aligned} \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} &= \sum_{j=1}^{m} a_{1j} + \sum_{j=1}^{m} a_{2j} + ... + \sum_{j=1}^{m} a_{nj} \\ &= a_{11} + a_{12} + a_{13} + ... + a_{1m} \\ &+ a_{21} + a_{22} + a_{23} + ... + a_{2m} \\ &+ \dots \\ &+ \dots \\ &+ \dots \\ &+ a_{n1} + a_{n2} + a_{n3} + ... + a_{nm} \end{aligned} \tag{29}$$

Consider the following matrix

$$A = \begin{pmatrix} 4 & -1 & 5 & 7 \\ 1 & 4 & -2 & 3 \\ 0 & 10 & -2 & 0 \\ -11 & 1 & 6 & 2 \end{pmatrix}$$

We compute $\sum_{i=2}^{4} \sum_{j=2}^{3} a_{ij}$ as follows

$$\begin{aligned} S &= a_{22} + a_{23} + a_{32} + a_{33} + a_{42} + a_{43} \\ &= 4 + -2 + 10 + -2 + 1 + 6 \\ &= 17 \end{aligned}$$

3.2.2. *Properties of a double sum.*

$$\begin{aligned} \left(\sum_{j=1}^{n} a_j\right) \left(\sum_{i=1}^{n} a_i\right) &= \sum_{i=1}^{n} a_i^2 + 2 \sum\sum_{i<j} a_i a_j \\ &= \sum_{i=1}^{n} a_i^2 + \sum\sum_{i \neq j} a_i a_j \end{aligned} \tag{30}$$

For example let a = $\begin{bmatrix} c & d & e \end{bmatrix}$. Then

$$[c + d + e] \times [c + d + e] = c^2 + d^2 + e^2 + 2cd + 2ce + 2de$$

4. Matrix operations

4.1. **Scalar multiplication (matrix).** Given a matrix A and a scalar λ, the product of λ and A, written λA, is defined to be

$$\lambda\, A \;=\; \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix} \tag{31}$$

4.2. **Scalar multiplication (vector).** Given a column vector $\vec{a}$ and a scalar λ, the product of λ and $\vec{a}$, written $\lambda\,\vec{a}$, is defined to be

$$\lambda\,\vec{a} \;=\; \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_m \end{pmatrix} \tag{32}$$

For the second column of a matrix we could write

$$\lambda \vec{a}_2 \;=\; \begin{pmatrix} \lambda a_{12} \\ \lambda a_{22} \\ \lambda a_{32} \\ \vdots \\ \lambda a_{m2} \end{pmatrix} \tag{33}$$

4.3. **Trace of a square matrix.** The trace of a matrix is the sum of the diagonal elements and is denoted tr A. Consider the matrix C below.

$$C \;=\; \begin{pmatrix} 3 & 1 & 5 & 1 \\ 1 & 4 & 10 & 1 \\ 5 & 10 & -2 & 15 \\ 1 & 1 & 15 & -2 \end{pmatrix} \tag{34}$$

The trace of C is [3 + 4 + -2 + -2] = 3.

4.4. **Addition of vectors.** - The sum of a vector a with m elements and a vector b having m elements is a vector c with m elements and whose elements are given by

$$c_j \;=\; a_j \;+\; b_j \,\forall\, j \tag{35}$$

This gives

$$\vec{c} \;=\; \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} \;=\; \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \;+\; \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \;=\; \begin{pmatrix} a_1 \;+\; b_1 \\ a_2 \;+\; b_2 \\ \vdots \\ a_m \;+\; b_m \end{pmatrix} \tag{36}$$

4.5. **Linear combinations of vectors.** If a and b are two n-vectors and s and t are two real numbers, ta + sb is said to be the linear combination of a and b. In symbols we write,

$$t\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} + s\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} t\,a_1 + s\,b_1 \\ t\,a_2 + s\,b_2 \\ \vdots \\ t\,a_m + s\,b_m \end{pmatrix} \tag{37}$$

4.5.1. *Example.* Let $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and let $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$. Let t = 2 and s =4. Then we obtain

$$2\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + 4\begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 2a_1 + 4b_1 \\ 2a_2 + 4b_2 \\ 2a_3 + 4b_3 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

where $\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$ represents the linear combination.

4.5.2. *Numerical Example.*

$$3\begin{pmatrix} 5 \\ -2 \\ 4 \end{pmatrix} + 2\begin{pmatrix} -3 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 15 + -6 \\ -6 + 8 \\ 12 + 2 \end{pmatrix} = \begin{pmatrix} 9 \\ 2 \\ 14 \end{pmatrix}$$

4.5.3. *Writing a system of equations as a linear combination of vectors.*

Consider three vectors, each with two elements. Call the vectors $\vec{a}_1$, $\vec{a}_2$ and $\vec{b}$. Call the elements of the first one a_{11} and a_{21}, the elements of the second one a_{12} and a_{22} and the elements of $\vec{b}$, b_1 and b_2. Now consider two scalars denoted x_1 and x_2. Now multiply $\vec{a}_1$ by x_1 and $\vec{a}_2$ by x_2 and add the products. We obtain

$$x_1\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}\,x_1 \\ a_{21}\,x_1 \end{pmatrix} + \begin{pmatrix} a_{12}\,x_2 \\ a_{22}\,x_2 \end{pmatrix} = \begin{pmatrix} a_{11}\,x_1 + a_{12}\,x_2 \\ a_{21}x_1 + a_{22}\,x_2 \end{pmatrix} \tag{38}$$

If set this expression equal to $\vec{b}$ we obtain

$$\begin{pmatrix} a_{11}\,x_1 + a_{12}\,x_2 \\ a_{21}x_1 + a_{22}\,x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{39}$$

which is a linear system of 2 equations in 2 unknowns. We can write a general system of m equations in n unknowns as

$$x_1\,\vec{a}_1 + x_2\,\vec{a}_2 + \ldots + x_n\,\vec{a}_n = \vec{b} \tag{40}$$

where x_i are a series of scalar unknowns and each a_j is a column of the A matrix of coefficients.

4.6. **Addition of matrices.** The sum C of a matrix A having m rows and n columns and a matrix B having m rows and n columns is a matrix having m rows and n columns whose elements are given by

$$c_{ij} = a_{ij} + b_{ij} \,\forall\, i,j \tag{41}$$

This gives

$$\begin{aligned} C &= \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} \\ &= \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix} \end{aligned} \tag{42}$$

For example

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

4.7. **Inner (dot) product of two vectors.** The inner (scalar or dot) product to two vectors u,v of length n is the scalar quantity denoted by

$$u \cdot v = \sum_{i=1}^{n} u_i v_i = u_1 v_1 + u_2 v_2 + ... + u_n v_n \tag{43}$$

It is easiest to see how to multiply two vectors if we write the first one as a row vector and the second one as a column vector. For example

$$\begin{bmatrix} 3 & 4 & 6 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} = 3 \times 2 + 4 \times 3 + 6 \times 5 = 6 + 12 + 30 = 48 \tag{44}$$

4.8. **Multiplication of a matrix and a column vector.** We can multiply a matrix and column vector if the matrix has the same number of columns as there are elements in the column vector. The result of this multiplication is a column vector with the same number of elements as the matrix has rows. The i^{th} element of the resulting column vector is obtained as the dot product of the i^{th} row of the matrix and the column vector. Specifically for an m×n matrix A and an n×1 column vector b,

$$c_i = \sum_{k=1}^{n} a_{ik}\, b_k,\ i = 1, \ldots, m. \tag{45}$$

For example

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & -4 \\ 5 & 2 & 1 \end{pmatrix} \begin{pmatrix} -2 \\ 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 18 \\ -4 \\ 1 \end{pmatrix}$$

Specifically 1 × -2 + 2 × 4 + 4 × 3 = 18 and so on.

Or consider the example

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

4.9. **Multiplication of a row vector and a matrix.** We can multiply a row vector and a matrix if the matrix has the same number of rows as there are elements in the row vector. The result of this multiplication is a row vector with the same number of elements as the matrix has columns. The i^{th} element of the resulting row vector is obtained as the dot product of the row vector and the i^{th} column of A. Specifically for an m×n matrix A and an m×1 column vector b,

$$c_i = \sum_{k=1}^{m} b_k \, a_{ki} \, , \; i = 1, \ldots, n. \tag{46}$$

For example

$$\begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 2 & -1 & -2 \\ 2 & 1 & 4 \\ -2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 5 & 18 \end{pmatrix}$$

Specifically 1 × 2 + 2 × 2 + 4 × -2 = -2 and so on.

Here is a second example.

$$\begin{pmatrix} 5 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & -2 \\ 2 & 1 & 4 \\ -2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 12 & -2 & 1 \end{pmatrix}$$

4.10. **Multiplication of matrices.** Given an m×n matrix A and an n×r matrix B, the product AB is defined to be an m × r matrix C, whose elements are computed from the elements of A,B according to

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}, i = 1, \ldots, m, j = 1, \ldots, r. \tag{47}$$

In other words to obtain the ij^{th} element of c we take the i^{th} row of A and j^{th} column of B and form the inner product.

4.10.1. *Example 1.* Consider multiplying the following two matrices A and B.

$$A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & -4 \\ 5 & 2 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 & -1 & -2 \\ 2 & 1 & 4 \\ -2 & 1 & 3 \end{pmatrix} \tag{48}$$

We obtain the first element of the product by multiplying the first row of A by the first column of B.

$$\begin{bmatrix} 1 & 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 2 \\ -2 \end{bmatrix} = c_{11} = \begin{pmatrix} -2 & & \\ & & \\ & & \end{pmatrix} \tag{49}$$

We obtain the second element of the first row of the product by multiplying the first row of A by the second column of B.

$$\begin{bmatrix} 1 & 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = c_{12} = \begin{pmatrix} & 5 & \\ & & \\ & & \end{pmatrix} \tag{50}$$

We obtain the third element of the first row of the product by multiplying the first row of A by the third column of B.

$$\begin{bmatrix} 1 & 2 & 4 \end{bmatrix} \cdot \begin{bmatrix} -2 \\ 4 \\ 3 \end{bmatrix} = c_{13} = \begin{pmatrix} & & 18 \\ & & \\ & & \end{pmatrix} \tag{51}$$

Combining operations for the first row of A and the matrix B we obtain

$$\begin{pmatrix} 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} 2 & -1 & -2 \\ 2 & 1 & 4 \\ -2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 5 & 18 \\ & & \\ & & \end{pmatrix}$$

Now consider the second row of A and the matrix B. Performing these operations we obtain

$$\begin{pmatrix} 2 & 3 & -4 \end{pmatrix} \begin{pmatrix} 2 & -1 & -2 \\ 2 & 1 & 4 \\ -2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 18 & -3 & -4 \end{pmatrix}$$

Completing the operations we obtain

$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & 3 & -4 \\ 5 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & -1 & -2 \\ 2 & 1 & 4 \\ -2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 5 & 18 \\ 18 & -3 & -4 \\ 12 & -2 & 1 \end{pmatrix}$$

4.10.2. *Example 2.* As a second example consider the matrices below

$$A = \begin{pmatrix} 3 & 4 & 7 \\ 2 & 5 & 2 \\ 1 & 0 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & 4 & 1 \end{pmatrix} \tag{52}$$

The element c_{11} comes from multiplying the first row of A with the first column of B as follows:

$$c_{11} = \begin{pmatrix} 3 & 4 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = 3 + 8 + 7 = 18 \tag{53}$$

Similarly the element c_{32} comes from multiplying the third row of A with the second column of B as follows:

$$c_{32} = \begin{pmatrix} 1 & 0 & 4 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} = 0 + 0 + 16 = 16 \tag{54}$$

Multiplying out the rest of the entries gives

$$C = \begin{pmatrix} 18 & 32 & 14 \\ 14 & 13 & 9 \\ 5 & 16 & 5 \end{pmatrix} \tag{55}$$

4.11. **Writing a system of equations as a matrix product.** Consider an m×n matrix, an n×1 vector and an m×1 vector. The case of a square matrix is handled by setting m=n. Call the matrix A and the vectors $\vec{x}$ and $\vec{b}$. Call the elements of $\vec{x}$, $x_1, x_2, \ldots, x_n$, and the elements of $\vec{b}$, $b_1, b_2, \ldots, b_n$. Consider a case where A is a 4×3 matrix and x is an 3×1 vector and b is a 4×1 vector. Multiply the matrix A by the column vector x and set it equal to the vector b as follows.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} \tag{56}$$

If we then carry out the multiplication we obtain

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 &= b_4 \end{aligned} \tag{57}$$

which is a linear system of 4 equations in 3 unknowns. The general system of m equations in n unknowns can be written

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & \cdots & + & a_{3n}x_n & = & b_3 \\ \vdots & + & \vdots & + & \cdots & + & \vdots & = & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \tag{58}$$

In this system, the a_{ij}'s and b_i's are given real numbers; a_{ij} is the coefficient for the unknown x_j in the i^{th} equation. We call the set of all a_{ij}'s arranged in a rectangular array the coefficient matrix of the system. Using matrix notation we can write the system as

$$Ax = b$$

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \tag{59}$$

Consider the following matrix A and vector b

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ -3 & -4 & -2 \end{pmatrix}, \qquad b = \begin{pmatrix} 3 \\ 8 \\ -4 \end{pmatrix} \tag{60}$$

We can then write

$$
\begin{aligned}
Ax &= b \\
\begin{pmatrix} 1 & 2 & 1 \\ 2 & 5 & 2 \\ -3 & -4 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= \begin{pmatrix} 3 \\ 8 \\ -4 \end{pmatrix}
\end{aligned} \tag{61}
$$

for the linear equation system

$$
\begin{aligned}
x_1 + 2x_2 + x_3 &= 3 \\
2x_1 + 5x_2 + 2x_3 &= 8 \\
-3x_1 - 4x_2 - 2x_3 &= -4
\end{aligned} \tag{62}
$$

4.12. **A system of equations with an identity coefficient matrix.** Consider a system of n variables and n equations. The coefficient matrix is square. If the coefficient matrix is an identity matrix then the solution is obvious upon inspection.

$$
\begin{aligned}
Ix &= b \\
\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} &= \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{pmatrix}
\end{aligned} \tag{63}
$$

Consider the following 3×3 example.

$$
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \\ 1 \end{pmatrix} \tag{64}
$$

It is obvious that x_1 = -2 and so on as $1 \times x_1 + 0 \times x_2 + 0 \times x_3 = -2$. If one were to use Gaussian elimination to solve a system of n equations in n unknowns and rewrite the system as a matrix equation at each step, it is clear one would end up with a system where the coefficient matrix was an identity matrix.

4.13. **A system of equations with a diagonal coefficient matrix.** Consider a system of n variables and n equations. The coefficient matrix is square. If the coefficient matrix is a diagonal matrix then the solution can be obtained by solving each equation individually by one simple division.

$$
\begin{aligned}
Ix &= b \\
\begin{pmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} &= \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{pmatrix}
\end{aligned} \tag{65}
$$

It is clear that $a_{11}\, x_1 = b_1$ which implies that $x_1 = \frac{b_1}{a_{11}}$. Similarly $x_2 = \frac{b_2}{a_{22}}$ and so on.
Consider the following 3×3 example.

$$\begin{pmatrix} -3 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \\ 5 \end{pmatrix} \tag{66}$$

It is obvious that (-3)x_1 = 6 which implies that x_1 = $\frac{6}{-3}$ = -2 and so on.

4.14. **Some properties of matrix operations.** Let α and β denote real numbers (scalars), $\vec{a}, \vec{b}, \vec{c}$ denote n-vectors, and A, B, C denote matrices. The properties are conditional on the operations being defined for the case in point.

4.14.1. *Equality.*

vectors: Two n-vectors a and b are said to be equal if all their corresponding components are equal. Equality is only possible for vectors of the same dimension.

matrices: Two m x n matrices A and B are said to be equal if all their corresponding components are equal. Equality is only possible for matrices of the same dimension.

4.14.2. *Multiplication by a scalar.*

a: (α + β) A = αA + βA
b: α(A + B) = αA + αB
c: α (βA) = ($\alpha\beta$) A

Note that A and B above can be replaced by a and b as in (1)(a) = a

4.14.3. *Addition.*

a: $\vec{a} + \vec{b} = \vec{b} + \vec{a}$
b: $\vec{a} + 0 = \vec{a}$
c: $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$
d: $\vec{a} + (-\vec{a}) = 0$
e: $A + B = B + A$
f: $A + (B + C) = (A + B) + C$
g: $A + 0 = 0 + A = A$
h: $A + (-A) = 0$

4.14.4. *Multiplication.*

a: $\vec{a}\,\vec{b} = \vec{b}\,\vec{a}$
b: AB $\neq$ BA
c: A(BC) = (AB)C
d: $\alpha(\vec{b} + \vec{c}) = \alpha\vec{b} + \alpha\vec{c}$
e: $A(B + C) = AB + AC$
f: $(B + C)A = AB + CA$
g: $(\alpha\vec{a})\vec{b} = \vec{a}(\alpha\vec{b}) = \alpha(\vec{a}\vec{b})$
h: $\vec{a} \cdot \vec{a} > 0 \Leftrightarrow \vec{a} \neq 0$
i: $\vec{a} \cdot 0 = 0 \cdot \vec{a} = 0$
j: $A0 = 0A = 0$
k: $AI = IA = A$

4.14.5. *Transposes.*

a: $(A')' = A$
b: $(ABC)' = C'\,B'\,A'$
c: $(A + B)' = A' + B'$

4.14.6. *Properties of the trace.*

- **a:** trace (I) = n
- **b:** trace (ABC) = trace (CAB) = trace (BCA)
- **c:** trace (A + B) = trace (A) + trace (B)
- **d:** tr(AB) = tr(BA) if both AB and BA are defined
- **e:** tr(kA) = ktr(A) where k is a scalar

4.15. **Idempotent matrices.** - A matrix is called idempotent if

$$A^2 = A \tag{67}$$

For example the identity matrix is idempotent. Consider the matrix M below.

$$M = \begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \\ -0.2 & 0.8 & -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & 0.8 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.8 & -0.2 \\ -0.2 & -0.2 & -0.2 & -0.2 & 0.8 \end{pmatrix} \tag{68}$$

We can verify that it is idempotent by carrying out the multiplication.

$$M\,M = \begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \\ -0.2 & 0.8 & -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & 0.8 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.8 & -0.2 \\ -0.2 & -0.2 & -0.2 & -0.2 & 0.8 \end{pmatrix} \begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \\ -0.2 & 0.8 & -0.2 & -0.2 & -0.2 \\ -0.2 & -0.2 & 0.8 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.8 & -0.2 \\ -0.2 & -0.2 & -0.2 & -0.2 & 0.8 \end{pmatrix} \tag{69}$$

Consider the multiplication of the first row and first column

$$\begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \end{pmatrix} \begin{pmatrix} 0.8 \\ -0.2 \\ -0.2 \\ -0.2 \\ -0.2 \end{pmatrix} = 0.64 + 0.4 + 0.4 + 0.4 + 0.4 = 0.8 \tag{70}$$

Or consider the multiplication of the first row and second column

$$\begin{pmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \end{pmatrix} \begin{pmatrix} -0.2 \\ 0.8 \\ -0.2 \\ -0.2 \\ -0.2 \end{pmatrix} = -0.16 + -0.16 + 0.4 + 0.4 + 0.4 = -0.2 \tag{71}$$

Later we will discuss an important concept called the rank of a matrix. For an idempotent matrix A, tr(A) = rank of A .