1 Operations with Matrices

$$\text{Denote matrix : } \begin{cases} A, B, C, \ldots \text{ or } [a_{ij}], [b_{ij}], [c_{ij}] \\ \text{A rectangular array of numbers } \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \end{cases}$$

Equality of Matrices.

$$A = [a_{ij}] = [b_{ij}] = B$$

for all $1 \le i \le m$ **and** $1 \le j \le n$

- Column matrix - Column vector $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$
- Row matrix - Row vector $\begin{bmatrix} 1 & 3 \end{bmatrix}$

Matrix Addition

$A + B = [a_{ij} + b_{ij}]$

Scalar Multiplication

$$cA = [ca_{ij}]$$

Matrix Multiplication

$A = [a_{ij}]$ is an $m \times n$ matrix and $B = [b_{ij}]$ is an $n \times p$ matrix, then the **product** AB is an $m \times p$ matrix : $AB = [c_{ij}]$
use the i^{th} row of A and the j^{th} column of B

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

- not **commutative** : AB is not always equal to BA

Systems of Linear Equations

A : the coefficient matrix of the system
x and **b** are column matrices.

$$\text{Rewrite the system as } \underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{23} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}}_{b}$$

Partitioned Matrix

$$b = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \sum_{j=1}^{n} x_j a_j$$

That is, a **linear combination** of the column matrices $a_1, a_2, \ldots, a_n$ with **coefficients** $x_1, x_2, \ldots, x_n$

Properties of Matrix Operations

1. $-A$: **additive inverse** of A
2. $A + B = B + A$ (Commutative property of addition)
3. $A + (B + C) = (A + B) + C$ (Associative property of addition)
4. $(cd)A = c(dA)$ (Associative property of multiplication)
5. $IA = A$ (Multiplicative identity)
6. $c(A + B) = cA + cB$ (Distributive property)
7. $(c + d)A = cA + dA$ (Distributive property)

Properties of Matrix Multiplication

1. not having general cancellation: $AC = BC$ doesn't mean $A = B$
2. AB is not always equal to BA
3. $(AB)C = A(BC)$
4. $A(B + C) = AB + AC$
5. $(A + B)C = AB + AC$
6. $c(AB) = A(cB)$

Identity Matrix of order n

$$I_n = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}$$

A is a $m \times n$ matrix : $AI_n = A$
$I_m A = A$

(

The Transpose of a Matrix)

The **transpose** of a matrix is formed by writing its columns as rows.

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \quad \text{then} \quad A^T = \begin{bmatrix} a_{11} & a_{21} & \ldots & a_{m1} \\ \vdots & & & \\ a_{1n} & a_{2n} & \ldots & a_{mn} \end{bmatrix}$$

1. $(A^T)^T = A$
2. $(A+B)^T = A^T + B^T$
3. $(ABC)^T = C^T.B^T.A^T$
4. $(AA^T) = (AA^T)^T = (A^T)^T.A^T$. So AA^T is symetric.

The Inverse of A Matrix

An $n \times n$ matrix is **invertible** (or **nonsingular**) if there exists an $n \times n$ matrix A^{-1} such that

$$AA^{-1} = A^{-1}A = I_n$$

- A^{-1} : the (multiplicative) **inverse** of A
- Else, A is **noninvertible** (or **singular**)
- A^{-1} is **unique**

If $AB = I_n$, it can be shown that $BA = I_n$ as well.

Finding the Inverse of a Matrix

Notice: The column vectors $a_1, a_2, \ldots, a_n$ have the same coefficient matrix which is A.

$$\left[\begin{array}{cc|cc} 1 & 4 & 1 & 0 \\ -1 & -3 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{cc|cc} 1 & 0 & -3 & -4 \\ 0 & 1 & 1 & 1 \end{array}\right]$$

$$[A \vdots I] \rightarrow [I \vdots A^{-1}]$$

If A can't be row reduced to I_n, then A is **noninvertible** (or **singular**)

For 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{then} \quad A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

The determinator $ad - bc$: the **determinant** of A.

1. $(A^{-1})^{-1} = A$
2. $(A^k)^{-1} = A^{-1}.A^{-1} \ldots A^{-1} = (A^{-1})^k$
3. $(cA)^{-1} = \frac{1}{c}A^{-1}, c \neq 0$
4. $(A^T)^{-1} = (A^{-1})^T$
5. If A and B are **invertible**, then AB is **invertible** : $(AB)^{-1} = B^{-1}A^{-1}$
   Thus, reverse the order of multiplication to find the inverse
   $$(A_1A_2 \ldots A_n)^{-1} = A_n^{-1}A_{n-1}^{-1} \ldots A_1^{-1}$$
6. If C **is invertible matrix**
   $$AC = BC \Leftrightarrow A = B$$
   $$CA = CB \Leftrightarrow A = B$$

Systems of Equations

For *square* systems (those having the same number of Eq. as variables), you can use the theorem below to determine whether the system has a **unique solution**.

If A is invertible,then the linear equation $Ax = b$ has an unique solution $x = A^{-1}b$

2 Elementary Matrices

1. **Elementary matrix** E : can be obtained from I_n by **1 elementary row operation**. If that same operation is performed on $A_{m\times n}$, the resulting matrix is given by the product EA.
2. Matrix B is **row-equivalent** to A if there exists a finite number of elementary matrices $E_1, E_2, \dots, E_k$ such that
$$B = E_1E_2 \dots E_kA$$
3. E^{-1} exist and is an **elementary matrix**. The **inverse** E^{-1} is simply found by reverse the operation to get E.

A is **invertible** $\Leftrightarrow A = E_1E_2 \dots E_k$

Consider the **homogeneous S. Eq** represented as $Ax = O$. Since A is invertible, it only has the **trivial solution**. But this implies that the **augmented matrix** $[A \vdots O]$ can be rewritten in the form $[I \vdots O]$ using $E_1, E_2, \dots, E_k$.

We now have $E_k \dots E_2E_1A = I$. It follows that $A = E_1^{-1}E_2^{-1} \dots E_k^{-1}$.

Equivalent Conditions

If A is an $n \times n$ matrix, the following statements are equivalent.

1. A is **invertible**
2. $Ax = b$ has a unique solution for every $n \times 1$ column matrix b.
3. $Ax = O$ has only the trivial solution.
4. A is **row-equivalent** to I_n
5. A can be written as the product of elementary matrices.

3 The LU-Factorization

At the heart of the most efficient and modern algorithms for solving linear systems, $Ax = b$ is the so-called LU-factorization, in which the square matrix A is expressed as a product $A = LU$.

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

L : lower triangular matrix U : upper triangular matrix

By writing $Ax = LUx$ and letting $Ux = y$, you can solve x for 2 stages:

1. Solve $Ly = b$ for y
2. Solve $Ux = y$ for x

Finding the LU-Factorizations of a Matrix

Find the LU-factorization of $A = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 2 & -10 & 2 \end{bmatrix}$

Matrix	*Elementary Row Operation*	*Elementary Matrix*
$\begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 0 & -4 & 2 \end{bmatrix}$	$R_3 + (-2)R_1 \to R_3$	$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$
$\begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 14 \end{bmatrix}$	$R_3 + (4)R_2 \to R_3$	$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}$

The matrix U on the left is upper triangular, and $E_2E_1A = U$, or $A = E_1^{-1}E_2^{-1}U$. Because the product of lower triangular matrices

$$E_1^{-1}E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -4 & 1 \end{bmatrix}$$

is again a lower triangular matrix L, the factorization $A = LU$ complete.

Note that the multipliers are negatives of the corresponding entries in L. If U is obtained from A using row operation of adding a multiple of 1 row to another row below, then L is a lower triangular matrix with 1's along the diagonal. Furthermore, the negative of each multiplier is the same position as that of the corresponding zero in U.

1. Write $y = Ux$ and solve $Ly = b$ for y.
2. Solve $Ux = y$ for x.

Those steps are just **forward-substitution** and **back-substitution**, since it's all triangular matrices.

4 Applications of Matrix Operations

4.1 Stochastic Matrices

Give a finite set of *states* $S_1, S_2, \ldots, S_n$.
For instance, residents of a city may live downtown or in the suburbs. Voters may vote Democrat, Republican, or for a third party. Soft drink consumers may buy Coca-Cola, Pepsi Cola, or another brand. •$0 \le p_{ij} \le 1$: the probability that a member will change from the jth state to the ith state.

$$P = \begin{array}{c} \overbrace{\begin{array}{cccc} S_1 & S_2 & \cdots & S_n \end{array}}^{\textit{From}} \\ \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \end{array} \begin{array}{c} S_1 \\ S_2 \\ \vdots \\ S_n \end{array} \; \textit{To}$$

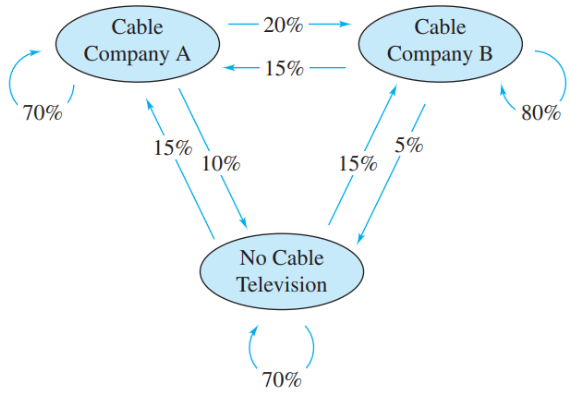
P: **matrix of transition probabilities**, since it gives the probabilities of each possible type of transition.

At each transition, each member in a given state $\begin{cases} \text{stay in that state} \\ \text{change to another state} \end{cases}$

This means the sum of the entries in any column of P is 1. For instance, in the first column

$$p_{11} + p_{21} + \cdots + p_{n1} = 1$$

In general, such a matrix is called **sochastic** (the term "sochastic" means "regarding conjecture").



The matrix representing the give transition probabilities is

$$P = \begin{array}{c} \overbrace{\begin{array}{ccc} \text{A} & \text{B} & \text{None} \end{array}}^{\textit{From}} \\ \begin{bmatrix} 0.70 & 0.15 & 0.15 \\ 0.20 & 0.80 & 0.15 \\ 0.10 & 0.05 & 0.70 \end{bmatrix} \end{array} \begin{array}{l} \left.\begin{array}{l} \text{A} \\ \text{B} \\ \text{None} \end{array}\right\} \textit{To} \end{array}$$

and the **state matrix** represent the current populations in 3 states is $X = \begin{bmatrix} 15,000 \\ 20,000 \\ 65,000 \end{bmatrix} \begin{array}{c} \text{A} \\ \text{B} \\ \text{None} \end{array}$

The state matrix after 1 year

$$PX = \begin{bmatrix} 0.70 & 0.15 & 0.15 \\ 0.20 & 0.80 & 0.15 \\ 0.10 & 0.05 & 0.70 \end{bmatrix} \begin{bmatrix} 15{,}000 \\ 20{,}000 \\ 65{,}000 \end{bmatrix} = \begin{bmatrix} 23{,}250 \\ 28{,}750 \\ 48{,}000 \end{bmatrix}.$$

After n years: $P^n X$. But when k reaches a specific value, the number of subcribers eventually reach a **steady state**. The product approaches a limit $\overline{X}$, that is $P\overline{X} = \overline{X}$.

4.2 Cryptography

A **cryptogram** is a message written according to a secret code (the Greek word *kryptos* means "hidden"). This section describes a method of using matrix multiplication to **encode** and **decode**.

Assign a number to each letter in the alphabet. The message is converted to numbers and partitioned into **uncoded row matrices**, each having n entries.

Forming Uncoded Row Matrices

Write the encoded row matrices of size 1×3 for the message MEET ME MONDAY.

Partitioning the message into into groups of 3.

$$\begin{array}{ccccccccccccccc} [13 & 5 & 5] & [20 & 0 & 13] & [5 & 0 & 13] & [15 & 14 & 4] & [1 & 25 & 0] \\ M & E & E & T & \_ & M & E & \_ & M & O & N & D & A & Y & \_ \end{array}$$

- **Encode**: Choose an $n \times n$ invertible matrix A and multiply the uncoded row matrices (gA).
- **Decode**: Multiply by A^{-1}. For those who do not know A, this is difficult.

4.3 Leontief Input-Output Models

Suppose that an economic system has n different industries $I_1, I_2, \ldots, I_n$, each of which has **input** needs (raw materials, utilities, etc.) and **output** (finished product). In producting each unit of output, an industry may use the outputs of others (including itself).

The input-output matrix :

$$D = \begin{array}{c} \overbrace{\begin{array}{cccc} I_1 & I_2 & \cdots & I_n \end{array}}^{\textit{User (Output)}} \\ \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix} \end{array} \left.\begin{array}{l} I_1 \\ I_2 \\ \vdots \\ I_n \end{array}\right\} \textit{Supplier (Input)}$$

where $0 \leq d_{ij} \leq 1$ is the amount of output of the jth industry needs from the ith one to produce 1 unit of output per year. Sum of all entries in each column does not exceed 1.

If the system is **closed**, let the total output of the ith industry be denoted by x_i.

$$x_i = d_{i1}x_1 + d_{i2}x_2 + \cdots + d_{in}x_n (\text{closed system})$$

On the other hand, if the industries within the system sell products to nonproducing groups (such as goverments or charitible organizations) outside the system, then the system is **open**

$$x_i = d_{i1}x_1 + d_{i2}x_2 + \cdots + d_{in}x_n + e_i (\text{closed system})$$

where e_i represents the external demand for the ith industry's product.

The matrix form of this system is

$$\boxed{X = DX + E}$$

where X is the **output matrix** and E is the **external demand matrix**.

$$(I - D)X = E$$

Applying Gauss-Jordan elimination on $(I - D)|E$ to solve for x.

4.4 Least Squares Regression Analysis

Develop Linear Models in Statistic.

One way of measuring how well $y = f(x)$ fit the given n points is to compute the differences between $f(x)$ and the actual y. (**sum of squared error**)

Of all possible linear models for a given data set, the one that has the best fit is defined to be the one that minimizes the sum of squared error. This models is called the **least squares regression line**, and the procedure for finding it is called the **method of least squares**.

For a set of points, the **least squares regression line** is given by the linear function

$$f(x) = a_0 + a_1 x$$

that minimizes the sum of squared error $[y_i - f(x_i)]^2$.

To find the least squares regression line for a set of points, begin by forming the system of linear equations

$$\begin{gathered} y_1 = f(x_1) + [y_1 - f(x_1)] \\ \vdots \\ y_n = f(x_n) + [y_n - f(x_n)] \end{gathered}$$

where the right hand term, $[y_i - f(x_i)]$ of each equation is thought of as the error in the approximation of y_i by $f(x_i)$. Write it as

$$e_i = y_i - f(x_i)$$

$$y_i = (a_o + a_i x_i) + e_i$$

Now define

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad A = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}, \quad E = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

the n linear equations may be replaced by the matrix equation

$$Y = XA + E$$

For the regression model $Y = XA + E$, the coefficients of the least squares regression line are given by the matrix equation

$$A = (X^T X)^{-1} X^T Y$$

and the sum of squared error is

$$E^T E$$