Linear Algebra

Chapter one

Matrices

A *matrix* is a rectangular array of numbers written between rectangular brackets, as in:

$$\begin{bmatrix} 1 & -2 & 0.5 & 2 \\ 2 & 0 & 6 & 1 \\ 1.5 & -1 & 2.1 & 0 \end{bmatrix}$$

It is also common to use large parentheses instead of rectangular brackets, as in:

$$\begin{pmatrix} 1 & -2 & 0.5 & 2 \\ 2 & 0 & 6 & 1 \\ 1.5 & -1 & 2.1 & 0 \end{pmatrix}$$

We represent the matrices as a capital letters, *A*,*B*,*C*,…, ete, and the elements of the matrix as a small letters , *a*,*b*,*c*,.. etc.

In general, With a real numbers *m*,*n*, a real-valued (*m*,*n*) **matrix** *A* is an $m{\cdot}n$-tuple of elements a_{ij}, *i = 1,2,…,m and j = 1,2,…,n*, which is ordered according to a rectangular scheme consisting of *m* rows and *n* columns:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad a_{ij} \in \mathbb{R}.$$

Therefore, the matrix *A* can be represents as $A = (a_{ij})_{mn}$ and the matrix *B* can be represents as : $B = (B_{ij})_{mxn}$ and so on.

An important attribute of a matrix is its *size* or *dimensions*, *i.e.*, the numbers of rows and columns. The matrix above has 3 rows and 4 columns, so its size is 3 ×4. A matrix of size *m* × *n* is called an *m* × *n* matrix.

The *elements* (or *entries* or *coefficients*) of a matrix are the values in the array. The *i*,*j* element is the value in the *i*th row and *j*th column, denoted by double subscripts: the *i*,*j* element of a matrix *A* is denoted a_{ij} (or $a_{i,j}$when *i* or *j* is more than one digit or character). The positive integers *i* and *j* are called the (row and column) *indices*. If *A* is an *m* × *n* matrix, then the row index *i* runs from 1 to *m* and the column index *j* runs from 1 to *n*. Row indices go from top to bottom, so row 1 is the top row and row *m* is the bottom row. Column indices go from left to right, so column 1 is the left column and column *n* is the right column.

If the matrix above is *B*, then we have$b_{13} = 0.5\,, b_{23} = 6$. The row index of the bottom left element (which has value 4*:*1) is 3; its column index is 1.

Two matrices are equal if they have the same size, and the corresponding entries are all equal.

Matrix indexing. As with vectors, standard mathematical notation indexes the rows and columns of a matrix starting from 1. In computer languages, matrices are often (but not always) stored as 2-dimensional arrays, which can be indexed in a variety of ways, depending on the language. Lower level languages typically use indices starting from 0; higher level languages and packages that support matrix operations usually use standard mathematical indexing, starting from 1.

Square, tall, and wide matrices. A *square* matrix has an equal number of rows and columns. A square matrix of size $n \times n$ is said to be of *order n*. A *tall* matrix has more rows than columns (size $m \times n$ with $m > n$). A *wide* matrix has more columns than rows (size $m \times n$ with $n > m$).

Column and row vectors. An *n*-vector can be interpreted as an $n \times 1$ matrix; we do not distinguish between vectors and matrices with one column. A matrix with only one row, *i.e.*, with size $1 \times n$, is called a *row vector*; to give its size, we can refer to it as an *n-row-vector*. As an example,

$$[\ -1.2 \quad -3 \quad 0]$$

is a 3-row-vector (or 1×3 matrix). To distinguish them from row vectors, vectors are sometimes called *column vectors*.

$$\begin{bmatrix} 1 \\ 2 \\ 1.5 \end{bmatrix}$$

A (1×1) matrix is considered to be the same as a scalar.

Columns and rows of a matrix. An $m{\times}n$ matrix *A* has *n* columns, given by (the *m*-vectors.

$$a_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}, for\ j = 1,2,\dots,n$$

The same matrix has *m* rows, given by the (*n*-row-vectors)

$$b_i = [a_{i1} \quad a_{i2} \cdots \quad a_{in}], for\ i = 1,2,\dots,n$$

As a specific example, the 2×3 matrix

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

has firs raw [1 2 3], (which is a 3-row-vector or a 1 × 3 matrix), and second column

$$\begin{bmatrix} 2 \\ 5 \end{bmatrix}$$

(which is a 2-vector or 2 × 1 matrix), also written compactly as $(2; 5)$.

Block matrices and submatrices. It is useful to consider matrices whose entries are themselves matrices, as in

$$A = \begin{bmatrix} B & C \\ D & E \end{bmatrix},$$

where B, C, D, and E are matrices. Such matrices are called *block matrices*; the elements B, C, D, and E are called *blocks* or *submatrices* of A. The submatrices can be referred to by their block row and column indices; for example, C is the *1,2* block of A.

Block matrices must have the right dimensions to fit together. Matrices in the same (block) row must have the same number of rows (*i.e.*, the same 'height'); matrices in the same (block) column must have the same number of columns (*i.e.*, the same 'width'). In the example above, B and C must have the same number of rows, and C and E must have the same number of columns. Matrix blocks placed next to each other in the same row are said to be *concatenated*; matrix blocks placed above each other are called ***stacked***.
As an example, consider:

$$B = [0 \quad 2 \quad 3], \;\; C[-1], D = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 5 \end{bmatrix}, \qquad E = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Then the block matrix A above is given by

$$A = \begin{bmatrix} 0 & 2 & 3 & -1 \\ 2 & 2 & 1 & 4 \\ 1 & 3 & 5 & 4 \end{bmatrix}$$

(Note that we have dropped the left and right brackets that delimit the blocks. This is similar to the way we drop the brackets in a 1 × 1 matrix to get a scalar).

We can also divide a larger matrix (or vector) into 'blocks'. In this context the blocks are called *submatrices* of the big matrix. As with vectors, we can use colon notation to denote submatrices. If A is an $m \times n$ matrix, and p, q, r, s are integers with $1 \leq p \leq q \leq m$ and $1 \leq r \leq s \leq n$, then $A_{p:q;r:s}$ denotes the submatrix

$$A_{p:q,r:s} = \begin{bmatrix} A_{pr} & A_{p,r+1} & \cdots & A_{ps} \\ A_{p+1,r} & A_{p+1,r+1} & & A_{P+1,s} \\ & \vdots & \ddots & \vdots \\ A_{qr} & A_{q,r+1} & \cdots & A_{qs} \end{bmatrix}$$

This submatrix has size (q -p + 1) ×(s-r + 1) and is obtained by extracting from A the elements in rows p through q and columns r through s.

For the specific matrix A above, we have

$$A_{2:3,3:4} = \begin{bmatrix} 1 & 4 \\ 5 & 4 \end{bmatrix}.$$

Column and row representation of a matrix. Using block matrix notation we can write an m × n matrix A as a block matrix with one block row and n block columns,

$$A = [a_1 \quad a_2 \ldots \quad a_n],$$

where a_j, which is an m-vector, is the jth column of A. Thus, an m × n matrix can be viewed as its n columns, concatenated.

Similarly, an m × n matrix A can be written as a block matrix with one block column and m block rows:

$$A = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

where b_i, which is a row n-vector, is the ith row of A. In this notation, the matrix A is interpreted as its m rows, stacked.

Examples:

Table interpretation. The most direct interpretation of a matrix is as a table of numbers that depend on two indices, i and j. (A vector is a list of numbers that depend on only one index.) In this case the rows and columns of the matrix usually have some simple interpretation. Some examples are given below.

- ***Images.*** A black and white image with M ×N pixels is naturally represented as an M × N matrix. The row index i gives the vertical position of the pixel, the column index *j* gives the horizontal position of the pixel, and the *i; j* entry gives the pixel value.

- ***Rainfall data.*** An *m* × *n* matrix *A* gives the rainfall at *m* different locations on *n* consecutive days, so A_{42} (which is a number) is the rainfall at location 4 on day 2. The *j*th column of *A*, which is an *m*-vector, gives the rainfall at the *m* locations on day *j*. The *i*th row of *A*, which is *n*-row-vector, is the time series of rainfall at location *i*.
- ***Asset returns.*** A *T* × *n* matrix *R* gives the returns of a collection of *n* assets (called the *universe* of assets) over *T* periods, with R_{ij} giving the return of asset *j* in period *i*. So $R_{12;7}$ = -0*:*03 means that asset 7 had a 3% loss in period 12. The 4th column of *R* is a *T*-vector that is the return time series for asset 4. The 3rd row of *R* in the *n*-row-vector that gives the returns of all assets in the universe in period 3.

  An example of an asset return matrix, with a universe of n = 4 assets over *T* = 3 periods, is shown in table 1 below:

  Table 1 Daily returns of Apple (AAPL), Google (GOOG), 3M (MMM), and Amazon (AMZN), on March 1, 2, and 3, 2016 (based on closing prices).

Date	AAPL	GOOG	MMM	AMZN
March 1, 2016	0.00219	0.00006	−0:00113	0.00202
March 2, 2016	0.00744	−0:00894	−0:00019	−0:00468
March 3, 2016	0.01488	−0:00215	0.00433	−0:00407

- ***Prices from multiple suppliers***. An *m* × *n* matrix *P* gives the prices of *n* different goods from *m* different suppliers (or locations): P_{ij} is the price that supplier *i* charges for good *j*. The *j*th column of *P* is the *m*-vector of supplier prices for good *j*; the *i*th row gives the prices for all goods from supplier *i*.

- ***Contingency table.*** Suppose we have a collection of objects with two attributes, the first attribute with *m* possible values and the second with *n* possible values. An *m* × *n* matrix *A* can be used to hold the counts of the numbers of objects with the different pairs of attributes: A_{ij} is the number of objects with first attribute *i* and second attribute *j*. (This is the analog of a count *n*-vector, that records the counts of one attribute in a collection.) For example, a population of college students can be described by a 4 × 50 matrix, with the *i, j* entry the number of students in year *i* of their studies, from state *j* (with the states ordered in, say, alphabetical order). The *i*th row of *A* gives the geographic distribution of students in year *i* of their studies; the *j*th column of *A* is a 4-vector giving the numbers of student from state *j* in their first through fourth years of study.

- ***Customer purchase history.*** An $n \times N$ matrix P can be used to store a set of N customers' purchase histories of n products, items, or services, over some period. The entry P_{ij} represents the dollar value of product i that customer j purchased over the period (or as an alternative, the number or quantity of the product). The jth column of P is the purchase history vector for customer j; the ith row gives the sales report for product i across the N customers.

Matrix representation of a relation or graph.

Suppose we have n objects labeled $1, 2, \dots, n$. A *relation* $\boldsymbol{\mathcal{R}}$ on the set of objects $\{1,2,\dots,n\}$ is a subset of ordered pairs of objects. As an example, R can represent a *preference relation* among n possible products or choices, with $(i, j) \in \boldsymbol{\mathcal{R}}$ meaning that choice i is preferred to choice j. A relation can also be viewed as a *directed graph*, with nodes (or vertices) labeled $1, 2, \dots, n$, and a directed edge from j to i for each $(i, j) \in \boldsymbol{\mathcal{R}}$. This is typically drawn as a graph, with arrows indicating the direction of the edge, as shown in figure 1, for the relation on 4 objects

$$\boldsymbol{\mathcal{R}} = \{(1,2), (1,3), (2,1), (2,4), (3,4), (4,1)\}$$

A relation $\boldsymbol{\mathcal{R}}$ on $\{1,2,\dots,n\}$ is represented as $nxn\ matrix\ A$ with

$$A_{ij} = \begin{cases} 1 & (i,j) \in \boldsymbol{\mathcal{R}} \\ 0 & (i,j) \notin \boldsymbol{\mathcal{R}} \end{cases} \qquad (1.1)$$

This matrix is called the ***adjacency matrix*** associated with the graph. (Some authors define the adjacency matrix in the reverse sense, with $A_{ij} = 1$ meaning there is an edge from i to j.) The relation above for example, is represented by the matrix

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

This is the adjacency matrix of the associated graph, shown in figure 1.

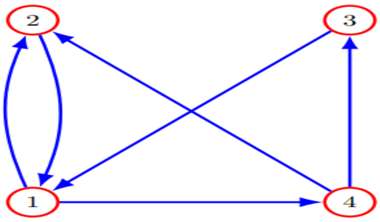


Figure 1 The relation (1.1) as a directed graph.

Types of matrices

Before we proceed to study the types of matrices we see it is better to mention some expressions that relate to matrices, as follows: $if\ A = (a_{ij})\ an\ mxn\ matrix, then$:

1. We say the entries $a_{ij}\ of\ A$ a main diagonal of *A if i=j.*
2. We say the entries $a_{ij}\ of\ A$ a secondary diagonal of *A if i+j=n+1.*

There are many types of matrices resulting from different application fields. In this lecture we focus on some of important matrices which are widely used in this course.

1. ***Square matrix***: Let $A = (a_{ij})$ be an mxn matrix, then A is said to be a ***square matrix*** if $m = n$.

   Examples: $A = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}, B = \begin{bmatrix} 1 & 5 & 1 \\ 2 & -1 & 4 \\ -2 & 0 & 2 \end{bmatrix}$

2. **Zero matrix**:$A\ matrix\ A = (a_{ij})\ is\ called\ a\ \boldsymbol{zero\ matrix}\ if\ a_{ij} = 0\ for\ all\ i\ and\ j.$

   Examples: $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

3. ***Identity matrix***: A square matrix $A = a_{ij}$ is called an ***identity matrix*** if it is satisfies the condition: $a_{ij} = \begin{cases} 1\ if\ i = j \\ 0\ if\ i \neq j \end{cases}$ and is denoted by I_n

   Examples: $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ldots, I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$

4. ***Diagonal matrix***: The identity matrix is said to be ***diagonal matrix*** if at least one of the entries of the main diagonal not equal one and is denoted by D_n.

   Examples: $D_2 = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}, D_3 = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \ldots, D_n = \begin{bmatrix} 0.5 & 0 & \cdots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$

   The notation **diag**(a_1, a_2,…,a_n) is used to compactly describe the $n \times n$ diagonal matrix A with diagonal entries $A_{11} = a_1, \ldots, A_{nn} = a_n$. As examples, the matrices above would be expressed as:
   diog(2,-1) , **diog**(3,1,-2), **diog**(0.5,1, …,1)

5. ***Triangular matrix***: A square matrix is called ***lower triangular*** if all the entries above the main diagonal are zero (or $a_{ij} = 0\ for\ all\ i > j$).

   Example: $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -9 & 0 \\ 4 & -3 & 2 \end{bmatrix}$

   Similarly, a square matrix is called ***upper triangular*** if all the entries below the main diagonal are zero (or $a_{ij} = 0\ for\ all\ i < j$).

Example: $A = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -9 & -2 \\ 0 & 0 & 2 \end{bmatrix}$

A ***triangular matrix*** is one that is either lower triangular or upper triangular. One can see that a matrix that is both upper and lower triangular is called a ***diagonal matrix***.

Remarks:

Two matrices *A* and *B* are said to be equal (written *A*=*B*) if and only if they have the same size and every entry in the matrix *A* is equal to the corresponding entry in *B*,

i.e, if

$$A = (a_{ij})_{mxn}, B = (b_{ij})_{pxk}, then\ A = B\ iff$$

i. $m = p\ and\ n = k$.
ii. $a_{ij} = b_{ij}\ for\ all\ i\ and\ j.$

Example: $Let\ A = \begin{bmatrix} 2 & 5 \\ -1 & 3 \end{bmatrix}, B = \begin{bmatrix} 2 & 5 \\ -1 & 3 \end{bmatrix}, then\ A = B.$

Matrix transpose

If *A* is an *m* × *n* matrix, its *transpose*, denoted A^T(or sometimes $A^t or\ A^*$) is the *n* × *m* matrix given by$(A^T)_{ij} = A_{ji}$. In words, the rows and columns of *A* are transposed in A^T . For example,

$$\text{if } \boldsymbol{A} = \begin{bmatrix} \mathbf{1} & \mathbf{4} \\ \mathbf{2} & \mathbf{5} \\ \mathbf{3} & \mathbf{6} \end{bmatrix}, \text{then } \boldsymbol{A^T} = \begin{bmatrix} \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{4} & \mathbf{5} & \mathbf{6} \end{bmatrix}$$

If we transpose a matrix twice, we get back the original matrix: $(\boldsymbol{A^T})^{\boldsymbol{T}} = \boldsymbol{A}.$

Transpose of block matrix. The transpose of a block matrix has the simple form (shown here for a 2 × 2 block matrix)

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix}^{\boldsymbol{T}} = \begin{bmatrix} \boldsymbol{A^T} & \boldsymbol{B^T} \\ \boldsymbol{C^T} & \boldsymbol{D^T} \end{bmatrix}$$

where *A*, *B*, *C*, and *D* are matrices with compatible sizes. The transpose of a block matrix is the transposed block matrix, with each element transposed.

Operations of Transposition

If A and B are $m \times n$ matrices, then the following hold:

$$(A+B)^T = A^T + B^T$$

$$(cA)^T = cA^T$$

$$(\boldsymbol{A^T})^{\boldsymbol{T}} = \boldsymbol{A}$$

Symmetric matrix.

A square matrix A is ***symmetric*** if $A = A^T, i.e., A_{ij} = A_{ji}\ for\ all\ i, j.$ Symmetric matrices arise in several applications. For example, suppose that A is the adjacency matrix of a graph or relation. The matrix A is symmetric when the relation is symmetric, *i.e.*, whenever $(i, j) \in \boldsymbol{\mathcal{R}}$, we also have $(j, i) \in \boldsymbol{\mathcal{R}}$. An example is the *friend relation* on a set of *n* people, where $(i, j) \in \boldsymbol{\mathcal{R}}$ means that person *i* and person *j* are friends. (In this case the associated graph is called the 'social network graph').

Example:

$$Let\ A = \begin{bmatrix} 2 & 5 \\ 5 & 3 \end{bmatrix}\ then\ A\ is\ symmetric\ matris\ since\ A^T = \begin{bmatrix} 2 & 5 \\ 5 & 3 \end{bmatrix}$$

Simple operation on matrices

Matrix Addition and Subtraction

The first matrix operations we discuss are matrix addition and subtraction. The rules for these operations are simple.

Two matrices can be added (or subtracted) if and only if they have the same dimensions.

To add (or subtract) two matrices of the same dimensions, we add (or subtract) the corresponding entries. More formally, if *A* and *B* are *m* × *n* matrices, then *A* + *B* and *A* - *B* are the *m* × *n* matrices whose entries are given by:

$(\boldsymbol{A}+\boldsymbol{B})_{\boldsymbol{ij}} = \boldsymbol{A}_{\boldsymbol{ij}} + \boldsymbol{B}_{\boldsymbol{ij}}$ ijth entry of the sum=sum of the *ij* th entry

$(\boldsymbol{A}-\boldsymbol{B})_{\boldsymbol{ij}} = \boldsymbol{A}_{\boldsymbol{ij}} - \boldsymbol{B}_{\boldsymbol{ij}}$ ijth entry of the difference=difference of the *ij* th entry

Examples:

1- $\begin{bmatrix} 2 & 3 \\ -1 & 5 \\ 4 & -2 \end{bmatrix} + \begin{bmatrix} 9 & -5 \\ 0 & 3 \\ -1 & -3 \end{bmatrix} = \begin{bmatrix} 11 & -2 \\ -1 & 8 \\ 3 & -5 \end{bmatrix}$

2- $\begin{bmatrix} 2 & 3 \\ -1 & 5 \\ 4 & -2 \end{bmatrix} - \begin{bmatrix} 9 & -5 \\ 0 & 3 \\ -1 & -3 \end{bmatrix} = \begin{bmatrix} -7 & 8 \\ -1 & 2 \\ 5 & 1 \end{bmatrix}$

Scalar multiplication

If A is an $m \times n$ matrix and c is a real number, then cA is the $m \times n$ matrix obtained by multiplying all the entries of A by c. (We usually use lowercase letters c, d, e, ... to denote scalars.) Thus, the $i\,j$th entry of cA is given by:

$$c(A)_{ij} = (cA)_{ij}$$

In words, this rule is: To get the $i\,j$th entry of cA, multiply the $i\,j$th entry of A by c.

Example:

1. $3\begin{bmatrix} 9 & -5 \\ 0 & 3 \\ -1 & -3 \end{bmatrix} = \begin{bmatrix} 27 & -15 \\ 0 & 9 \\ -3 & -9 \end{bmatrix}$
2. $0.5\begin{bmatrix} 9 & -5 \\ 0 & 3 \\ -1 & -3 \end{bmatrix} = \begin{bmatrix} 4.5 & -2.5 \\ 0 & 1.5 \\ -0.5 & -1.5 \end{bmatrix}$

Properties of matrix addition and scalar multiplication

If A, B, and C are any $m \times n$ matrices and if O is the zero $m \times n$ matrix, then the following hold:

$A + (B + C) = (A + B) + C$	*Associative law*
$A + B = B + A$	*Commutative law*
$A + O = O + A = A$	*Additive identity law*
$A + (-A) = O = (-A) + A$	*Additive inverse law*
$c(A + B) = cA + cB$	*Distributive law*
$(c + d)A = cA + dA$	*Distributive law*
$1A = A$	*Scalar unit*
$0A = O$	*Scalar zero*

Matrix Multiplication

It is possible to multiply two matrices using *matrix multiplication*. You can multiply two matrices A and B provided their dimensions are *compatible*, which means the number of columns of A equals the number of rows of B. Suppose A and B are compatible, *e.g.*, A has size $m \times p$ and B has size $p \times n$. Then the product matrix $C = AB$ is the $m \times n$ matrix with elements

$$C_{ij} = \sum_{k=1}^{p} A_{ik}B_{kj} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{ip}B_{pj}\ ,\ \ i = 1,2, \ldots m, j = 1,2, \ldots, n.$$

There are several ways to remember this rule. To find the i; j element of the product C = AB, you need to know the ith row of A and the jth column of B. The summation above can be interpreted as ‘moving left to right along the ith row of A’ while moving ‘top to bottom’ down the jth column of B. As you go, you keep a running sum of the product of elements, one from A and one from B. As a specific example, we have

$$\begin{bmatrix} -1.5 & 3 & 2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} -1 & -1 \\ 0 & -2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 3.5 & -4.5 \\ -1 & 1 \end{bmatrix}$$

To find the 1*;* 2 entry of the right-hand matrix, we move along the first row of the left-hand matrix, and down the second column of the middle matrix, to get (-1.5)(-1) + (3)(-2) + (2)(0) = -4.5.

Remarks:

1. $AI=IA=A$
2. $AB \neq BA$ in general
3. If $AB = BA$ then A and B are said ***commute***. (Note that for $AB = BA$ to make sense, A and B must both be square.)

Properties of matrix multiplication. The following properties hold and are easy to verify from the definition of matrix multiplication. We assume that A, B, and C are matrices for which all the operations below are valid, and that γ is a scalar.

- *Associativity:* $(AB)C = A(BC)$.
- *Associativity with scalar multiplication:* $\gamma(AB) = (\gamma A)B$, where γ is a scalar.
- *Distributivity with addition:* $A(B+C) = AB+AC$ and $(A+B)C = AC+BC$
- *Transpose of product:* $(AB)^T = B^T A^T$
- $(A + B)(C + D) = AC + AD + BC + BD$

Products of block matrices. Suppose A is a block matrix with $m \times p$ block entries A_{ij}, and B is a block matrix with $p \times n$ block entries B_{ij}, and for each $k = 1, 2, \ldots, p$, the matrix product $A_{ik}B_{kj}$ makes sense, *i.e.*, the number of columns of A_{ik} equals the number of rows of B_{kj}. (In this case we say that the block matrices ***conform*** or are ***compatible***.) Then $C = AB$ can be expressed as the $m \times n$ block matrix with entries C_{ij}. For example, we have

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{bmatrix}$$

for any matrices A, B,C,D,E,F,G, H for which the matrix products above make sense. This formula is the same as the formula for multiplying two 2 × 2 matrices (i.e., with scalar entries); but when the entries of the matrix are themselves matrices (as in the block matrix above), we must be careful to preserve the multiplication order.

MATRIX DETERMINANTS

Summary

Uses

The determinant will be an essential tool to identify the maximum and minimum points or the saddle points of a function with multiple variables.

1- Reminder - Definition and components of a matrix

A matrix is a rectangular table of form

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

A matrix is said to be of dimension $m \times n$ when it has m rows and n columns. This method of describing the size of a matrix is necessary in order to avoid all confusion between two matrices containing the same amount of entries. For example, a matrix of dimension 3×4 has 3 rows and 4 columns. It would be distinct from a matrix 4×3, that has 4 rows and 3 columns, even if it also has 12 entries. A matrix is said to be **square** when it has the same number of rows and columns.

The elements are matrix entries a_{ij}, that are identified by their position. The element a_{32} would be the entry located on the third row and the second column of matrix A. This notation is essential in order to distinguish the elements of the matrix. The element a_{23}, distinct from a_{32}, is situated on the second row and the third column of the matrix A.

2- The matrix determinant

A value called the determinant of A, that we denote by

$$det(A) \text{ or } |A|,$$

corresponds to every square matrix A. We will avoid the formal definition of the determinant (that implies notions of permutations) for now and we will concentrate instead on its calculation.

3- Calculation of the determinant for a 2×2 matrix

Let us consider the matrix A of dimension 2×2 :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

The determinant of the matrix A is defined by the relation

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

The result is obtained by multiplying opposite elements and by calculating the difference between these two products…. a recipe that you will need to remember!

Example

Given the matrix

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

The determinant of A is

$$\det(A) = \begin{vmatrix} 2 & 1 \\ 3 & -2 \end{vmatrix}$$

4- Exercise

Calculate the determinant of the following 2×2 matrices :

$$a. \begin{pmatrix} 1 & 3 \\ 5 & -2 \end{pmatrix} \qquad b. \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}$$

$$c. \begin{pmatrix} 4 & -1 \\ -3 & 2 \end{pmatrix} \qquad d. \begin{pmatrix} 4 & -3 \\ 1 & 2 \end{pmatrix}$$

Solutions :   a) -17   b) 0   c) 5   d) 11

Before being able to evaluate the determinant of a 3×3 matrix (or all other matrices of a greater dimension), you will first need to learn a few concepts…

5- Definition of a minor

$$A = \begin{pmatrix} 2 & 1 & 4 \\ 5 & 2 & 3 \\ 8 & 7 & 3 \end{pmatrix}$$

The minor M_{12} is the determinant of the matrix obtained by eliminating the first row and the second column of A, i.e.

$$M_{12} = \begin{vmatrix} 5 & 3 \\ 8 & 3 \end{vmatrix} = 5.3 - 3.8 = 15 - 24 = -9$$

The minor M_{22} is the determinant of the matrix obtained by eliminating the second row and the second column of A, i.e.

$$M_{22} = \begin{vmatrix} 2 & 4 \\ 8 & 3 \end{vmatrix} = 2.3 - 4.8 = 6 - 32 = -26$$

6- Definition of a cofactor

The cofactor, C_{ij}, of a matrix A is defined by the relation

$$C_{ij} = (-1)^{i+j} M_{ij}$$

You will notice that the cofactor and the minor always have the same numerical value, with the possible exception of their sign.

Let us again consider the matrix

$$A = \begin{pmatrix} 2 & 1 & 4 \\ 5 & 2 & 3 \\ 8 & 7 & 3 \end{pmatrix}$$

We have already shown that the minor $M_{12} = -9$. Thus the corresponding cofactor, C_{12}, is

$$C_{12} = (-1)^{1+2} M_{12} = -1.(-9) = 9$$

The minor M_{12} and the cofactor C_{12} are of different signs.

The minor $M_{22} = -26$. Its corresponding cofactor C_{22} is

$$\mathrm{C}_{22} = (-1)^{2+2} \mathrm{M}_{22} = 1.(-26) = -26$$

This time, the minor M_{22} and the cofactor C_{22} are identical.

Evaluating the determinant of a 3×3 matrix is now possible. We will proceed by reducing it in a series of 2×2 determinants, for which the calculation is much easier. This process is called an **cofactor expansion.**

7- Cofactor expansion – a method to calculate the determinant

Given a square matrix A and its cofactors C_{ij}. The determinant is obtained by cofactor expansion as follows:

- Choose a row or a column of A (if possible, it is faster to choose the row or column containing the most zeros)…
- Multiply each of the elements a_{ij} of the row (or column) chosen by its corresponding cofactor, C_{ij}…
- Add these results.

8- Calculate the determinant for a 3×3 matrix

For a 3×3 matrix, this would mean that by choosing to make an expansion along the first row, the determinant would be

$$det\ A = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13}$$

If we had chosen to carry out an expansion along the second column, we would have to calculate

$$det\ A = a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32}$$

While the choice of row or column may differ, the result of the determinant will be the same, no matter what the choice we have made. Let us verify this with an example.

Example

What is the determinant of matrix A?

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 0 & 2 \\ 2 & 0 & -2 \end{pmatrix}$$

Solution

Let us follow the procedure proposed above (cofactor expansion):

- Choose a row or a column of A... For now, let us choose the first row.
- Multiply each of the elements of this row by their corresponding cofactors... The elements of the first row are $a_{11} = 2, a_{12} = 1,$ et $a_{13} = 3$ that we multiply with the corresponding cofactors, i.e. $\mathrm{C}_{11}, \mathrm{C}_{12}$ et C_{13}. These are

$$C_{11} = (-1)^{1+1}M_{11} = 1\begin{vmatrix} 0 & 2 \\ 0 & -2 \end{vmatrix} = 1(0.(-2) - 2.0) = 0$$

$$C_{12} = (-1)^{1+2}M_{12} = (-1)\begin{vmatrix} 1 & 2 \\ 2 & -2 \end{vmatrix} = 1(1 \times (-2) - 2 \times 2) = 6$$

$$\mathrm{C}_{13} = (-1)^{1+3}\mathrm{M}_{13} = 1\begin{vmatrix} 1 & 0 \\ 2 & 0 \end{vmatrix} = 1(1 \times (0) - 2 \times 0) = 0$$

Finally, we need to calculate

$$det\ A = a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32}$$

$$det\ A = 2 \times 0 + 6 \times 1 + 3 \times 0 = 6$$

Let us verify if an expansion along the second column coincides with the previous result. Note that the choice of the second column is much more effective since the determinant will be obtained from the calculation

$$\boldsymbol{det\ A = a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32}}$$

Two of the three elements of the second column are zero. In effect, $a_{12} = 1, a_{22} = 0, and\ a_{32} = 0$. It is thus useless to calculate the cofactors C_{22} and C_{32}. The corresponding cofactor for a_{12} is

$$\mathrm{C}_{12} = (-1)^{1+2}\mathrm{M}_{12} = (-1)\begin{vmatrix} 1 & 2 \\ 2 & -2 \end{vmatrix} = 1(1 \times (-2) - 2 \times 2) = 6$$

The determinant of A is thus

$$\det \mathrm{A} = \mathrm{a}_{12}\mathrm{C}_{12} + \mathrm{a}_{22}\mathrm{C}_{22} + \mathrm{a}_{32}\mathrm{C}_{32} = 1 \times 6 + 0 \times \mathrm{C}_{22} + 0 \times \mathrm{C}_{32} = 6,$$

which corresponds to the answer obtained by an expansion along the first row.

9- Alternative method to calculate determinants

This second method is in all points equivalent to cofactor expansion but will allow you to avoid the use of cofactors.

- Allocate a sign $+/-$ to each element by following the rule: we associate a positive sign to the position a_{11} , then we alternate the signs by moving horizontally or vertically.
- Choose a row or column of A (if possible, it is faster to choose the row or column of A containing the most number of zeros)…
- Multiply each element of a_{ij} of the row (or column) chosen by its corresponding minor, i.e. the remaining determinant when we eliminate the row and column in which a_{ij} is.
- Add or subtract these results according to the sign allocated to the elements during the first step.

Let us verify that this method will produce the same result as in the previous example:

Example

Given the matrix A to which we allocate a sign $+/-$ according to the rule stated above.

$$A = \begin{pmatrix} 2^+ & 1^- & 3^+ \\ 1^- & 0^+ & 2^- \\ 2^+ & 0^- & -2^+ \end{pmatrix}$$

- Let us choose the third column (it is certainly not the best choice since the second row has the most zeros, but…)
- We then multiply each element by its corresponding minor:

$$3\begin{vmatrix} 1 & 0 \\ 2 & 0 \end{vmatrix} = 3 \times 0 = 0$$

$$2\begin{vmatrix} 2 & 1 \\ 2 & 2 \end{vmatrix} = 2 \times (-2) = -4$$

$$-2\begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix} = -2 \times (-1) = 2$$

- Finally, the respective signs of the elements of the third column tell us the operations to carry out between these values to obtain the determinant:

$$\det A = \color{red}{+}\,0 \,\color{red}{-}\, (-4) \,\color{red}{+}\, 2 \,=\, 6$$

10- Exercise

Calculate the determinant of the following matrices:

$$a)\begin{pmatrix} 1 & 3 & 2 \\ 4 & 1 & 3 \\ 2 & 2 & 0 \end{pmatrix} \qquad b)\begin{pmatrix} 1 & 0 & 2 \\ 1 & 3 & 4 \\ 0 & 6 & 0 \end{pmatrix}$$

$$c)\begin{pmatrix} 3 & -2 & 4 \\ 2 & -4 & 5 \\ 1 & 8 & 2 \end{pmatrix} \qquad d)\begin{pmatrix} 8 & -1 & 9 \\ 3 & 1 & 8 \\ 11 & 0 & 17 \end{pmatrix}$$

Solutions : a) 24 b) -12 c) -66 d) 0

11- Determinants of square matrices of dimensions 4x4 and greater

The methods presented for the case of 3×3 matrices remain valid for all greater dimensions. You must again follow the steps for cofactor expansion:

Given a square matrix A and its cofactors C_{ij}, the determinant is obtained by following a cofactor expansion as follows:

- Chose a row or column of A (if possible, it is faster to choose the row or column that contains the most zeros) …
- Multiply each of the elements a_{ij} of the row (or column) chosen, by the corresponding cofactor C_{ij}…
- Add the results.

We must however mention a distinction. The cofactor associated to the element a_{ij} of a 4×4 matrix is the determinant of a 3×3 matrix, since it is obtained by eliminating the ith row and the jth column of A.

Example

Calculate the determinant of matrix A

$$A = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 3 & 1 & 1 \\ -1 & 0 & 3 & 1 \\ 3 & 1 & 2 & 0 \end{pmatrix}$$

It is essential, to reduce the amount of calculations, to choose the row or column that contains the most zeros (here, the fourth column). We will proceed to a cofactor expansion along the fourth column, which means that

$$det\ A = a_{14}C_{14} + a_{24}C_{24} + a_{34}C_{34} + a_{44}C_{44}$$

As a_{14} and a_{44} are zero, it is useless to find C_{14} and C_{44}. The cofactors C_{24} and C_{34} will be necessary…

$$C_{24} = (-1)^{2+4}M_{24} = 1\begin{vmatrix} 1 & 2 & 1 \\ -1 & 0 & 3 \\ 3 & 1 & 2 \end{vmatrix}$$

$$C_{34} = (-1)^{3+4} M_{34} = -1 \begin{vmatrix} 1 & 2 & 1 \\ 0 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix}$$

We let the reader verify that $C_{24} = 18$ et $C_{34} = -2$. Consequently, the determinant of A is

$$det\ A = a_{14}C_{14} + a_{24}C_{24} + a_{34}C_{34} + a_{44}C_{44}$$

$$det\ A = 0 \times C_{14} + 1 \times 18 + 1 \times (-2) + 0 \times C_{44} = 16$$

Exercise

Show that the determinant of A in the previous example is 16 by a cofactor expansion along

a) The first row
b) The third column

The inverse of a matrix

Introduction

In this leaflet we explain what is meant by an inverse matrix and how it is calculated.

1. The inverse of a matrix

The **inverse** of a square $n \times n$ matrix A, is another $n \times n$ matrix denoted by A^{-1} such that

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

where I is the $n \times n$ identity matrix. That is, multiplying a matrix by its inverse produces an identity matrix. Not all square matrices have an inverse matrix. If the determinant of the matrix is zero, then it will not have an inverse, and the matrix is said to be **singular**. Only non-singular matrices have inverses.

2. A formula for finding the inverse

Given any non-singular matrix A, its inverse can be found from the formula

$$A^{-1} = \frac{\mathrm{adj}\, A}{|A|}$$

where $\mathrm{adj}\, A$ is the **adjoint matrix** and $|A|$ is the determinant of A. The procedure for finding the adjoint matrix is given below.

3. Finding the adjoint matrix

The adjoint of a matrix A is found in stages:

(1) Find the transpose of A, which is denoted by A^T. The transpose is found by interchanging the rows and columns of A. So, for example, the first column of A is the first row of the transposed matrix; the second column of A is the second row of the transposed matrix, and so on.

(2) The **minor** of any element is found by covering up the elements in its row and column and finding the determinant of the remaining matrix. By replacing each element of A^T by its minor, we can write down a *matrix of minors* of A^T.

(3) The **cofactor** of any element is found by taking its minor and imposing a **place sign** according to the following rule

$$\begin{pmatrix} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \dots & \dots & \dots & \ddots \end{pmatrix}$$

This means, for example, that to find the cofactor of an element in the first row, second column, the sign of the minor is changed. On the other hand to find the cofactor of an element in the second row, second column, the sign of the minor is unaltered. This is equivalent to multiplying the minor by '+1' or '−1' depending upon its position. In this way we can form a *matrix of cofactors* of A^T. This matrix is called the **adjoint** of A, denoted $\mathrm{adj}\, A$.

The matrix of cofactors of the transpose of A, is called the adjoint matrix, $\mathrm{adj}\, A$

This procedure may seem rather cumbersome, so it is illustrated now by means of an example.

Example

Find the adjoint, and hence the inverse, of $A = \begin{pmatrix} 1 & -2 & 0 \\ 3 & 1 & 5 \\ -1 & 2 & 3 \end{pmatrix}$.

Solution

Follow the stages outlined above. First find the transpose of A by taking the first column of A to be the first row of A^T, and so on:

$$A^T = \begin{pmatrix} 1 & 3 & -1 \\ -2 & 1 & 2 \\ 0 & 5 & 3 \end{pmatrix}$$

Now find the minor of each element in A^T. The minor of the element '1' in the first row, first column, is obtained by covering up the elements in its row and column to give $\begin{pmatrix} 1 & 2 \\ 5 & 3 \end{pmatrix}$ and finding the determinant of this, which is -7. The minor of the element '3' in the second column of the first row is found by covering up elements in its row and column to give $\begin{pmatrix} -2 & 2 \\ 0 & 3 \end{pmatrix}$ which has determinant -6. We continue in this fashion and form a new matrix by replacing every element of A^T by its minor. Check for yourself that this process gives

$$\text{matrix of minors of } A^T = \begin{pmatrix} -7 & -6 & -10 \\ 14 & 3 & 5 \\ 7 & 0 & 7 \end{pmatrix}$$

Then impose the place sign. This results in the matrix of cofactors, that is, the adjoint of A.

$$\mathrm{adj}\, A = \begin{pmatrix} -7 & 6 & -10 \\ -14 & 3 & -5 \\ 7 & 0 & 7 \end{pmatrix}$$

Notice that to complete this last stage, each element in the matrix of minors has been multiplied by 1 or -1 according to its position.

It is a straightforward matter to show that the determinant of A is 21. Finally

$$A^{-1} = \frac{\mathrm{adj}\, A}{|A|} = \frac{1}{21} \begin{pmatrix} -7 & 6 & -10 \\ -14 & 3 & -5 \\ 7 & 0 & 7 \end{pmatrix}$$

Exercise

1. Show that the inverse of $\begin{pmatrix} 1 & 3 & 2 \\ 0 & 5 & 1 \\ -1 & 3 & 0 \end{pmatrix}$ is $\frac{1}{4} \begin{pmatrix} -3 & 6 & -7 \\ -1 & 2 & -1 \\ 5 & -6 & 5 \end{pmatrix}$.

Elementary Row Operations

In matrices we are allowed to perform operations of the following types:

1. Interchange two rows in the matrix (ex. $Ri \leftrightarrow Rj$).
2. Multiply a row by a non-zero constant (ex. $Ri \rightarrow kRi$, where k is constant).
3. Add a multiple of one row to another row (ex. $Ri \rightarrow Ri + kRj\ , where\ k\ is\ constant$).

The above three operations are called ***elementary raw operation (ERO's)*** on a matrix. Note that we can perform these operations on columns of the matrices and in this case they called ***elementary column operations*** on a matrix.

Example:

The following table describes how an ERO is performed at each step to produce a new simpler matrix

$$A = \begin{bmatrix} 3 & 1 & -2 & 9 \\ 1 & 2 & -1 & 5 \\ -1 & 4 & 2 & 0 \end{bmatrix}$$

$$R1 \leftrightarrow R2 \qquad \begin{bmatrix} 1 & 2 & -1 & 5 \\ 3 & 1 & -2 & 9 \\ -1 & 4 & 2 & 0 \end{bmatrix}$$

$$R3 \rightarrow R3 + R1 \qquad \begin{pmatrix} 1 & 2 & -1 & 5 \\ 3 & 1 & -2 & 9 \\ 0 & 6 & 1 & 5 \end{pmatrix}$$

$$R2 \rightarrow R2 - 3R1 \qquad \begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & -5 & 1 & -6 \\ 0 & 6 & 1 & 5 \end{pmatrix}$$

$$R2 \rightarrow R2 + R3 \qquad \begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 6 & 1 & 5 \end{pmatrix}$$

$$R3 \rightarrow R3 - 6R2 \qquad \begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -11 & 11 \end{pmatrix}$$

$$R3 \times (-\tfrac{1}{11}) \qquad \begin{pmatrix} 1 & 2 & -1 & 5 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Using row reduction to calculate the inverse and the determinant of a square matrix

Hayder Kadhim Zghair

1 Inverse of a square matrix

An $n \times n$ square matrix $\mathbf{A}$ is called *invertible* if there exists a matrix $\mathbf{X}$ such that

$$\mathbf{AX} = \mathbf{XA} = \mathbf{I},$$

where $\mathbf{I}$ is the $n \times n$ identity matrix. If such matrix $\mathbf{X}$ exists, one can show that it is *unique.* We call it the *inverse of* $\mathbf{A}$ and denote it by $\mathbf{A^{-1}} = \mathbf{X}$, so that

$$\mathbf{AA^{-1}} = \mathbf{A^{-1}A} = \mathbf{I}$$

holds if $\mathbf{A^{-1}}$ exists, i.e. if $\mathbf{A}$ is invertible. *Not all matrices are invertible.* If $\mathbf{A^{-1}}$ does not exist, the matrix $\mathbf{A}$ is called *singular* or *noninvertible.*

Note that if $\mathbf{A}$ is invertible, then the linear algebraic system

$$\mathbf{Ax} = \mathbf{b}$$

has a *unique* solution $\mathbf{x} = \mathbf{A^{-1}b}$. Indeed, multiplying both sides of $\mathbf{Ax} = \mathbf{b}$ on the left by $\mathbf{A^{-1}}$, we obtain

$$\mathbf{A^{-1}Ax} = \mathbf{A^{-1}b}.$$

But $\mathbf{A^{-1}A} = \mathbf{I}$ and $\mathbf{Ix} = \mathbf{x}$, so

$$\mathbf{x} = \mathbf{A^{-1}b}$$

The converse is also true, so for a square matrix $\mathbf{A}$,

$\mathbf{Ax} = \mathbf{b}$ *has a unique solution if and only if* $\mathbf{A}$ *is invertible.*

2 Calculating the inverse

To compute $\mathbf{A^{-1}}$ if it exists, we need to find a matrix $\mathbf{X}$ such that

$$\mathbf{AX} = \mathbf{I} \tag{1}$$

Linear algebra tells us that if such $\mathbf{X}$ exists, then $\mathbf{XA} = \mathbf{I}$ holds as well, and so $\mathbf{X} = \mathbf{A^{-1}}$.

Now observe that solving (1) is equivalent to solving the following linear systems:

$$\begin{aligned}\mathbf{Ax_1} &= \mathbf{e_1}\\ \mathbf{Ax_2} &= \mathbf{e_2}\\ &\dots\\ \mathbf{Ax_n} &= \mathbf{e_n},\end{aligned}$$

where $\mathbf{x_j}$, $j = 1, \dots, n$, is the (unknown) jth column of $\mathbf{X}$ and $\mathbf{e_j}$ is the jth column of the identity matrix $\mathbf{I}$. If there is a unique solution for each $\mathbf{x_j}$, we can obtain it by using elementary row operations to reduce the augmented matrix $[\,\mathbf{A} \mid \mathbf{e_j}\,]$ as follows:

$$[\,\mathbf{A} \mid \mathbf{e_j}\,] \longrightarrow [\,\mathbf{I} \mid \mathbf{x_j}\,].$$

Instead of doing this for each j, we can row reduce all these systems simultaneously, by attaching *all* columns of $\mathbf{I}$ (i.e. the whole matrix $\mathbf{I}$) on the right of $\mathbf{A}$ in the augmented matrix and obtaining *all* columns of $\mathbf{X}$ (i.e. the whole inverse matrix) on the right of the identity matrix in the row-equivalent matrix:

$$[\,\mathbf{A} \mid \mathbf{I}\,] \longrightarrow [\,\mathbf{I} \mid \mathbf{X}\,].$$

If this procedure works out, i.e. if we are able to convert $\mathbf{A}$ to identity using row operations, then $\mathbf{A}$ is invertible and $\mathbf{A^{-1}} = \mathbf{X}$. If we *cannot* obtain the identity matrix on the left, i.e. we get a row of zeroes, then $\mathbf{A^{-1}}$ does *not* exist and $\mathbf{A}$ is *singular.*

Example 1. Find the inverse of

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3\\ 2 & 4 & 5\\ 3 & 5 & 6 \end{bmatrix}$$

or show that it does not exist.

Solution:

$$\text{form the augmented matrix } [\,\mathbf{A} \mid \mathbf{I}\,]\text{:} \quad \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0\\ 2 & 4 & 5 & 0 & 1 & 0\\ 3 & 5 & 6 & 0 & 0 & 1 \end{array}\right]$$

$$R_3 - 3R_1 \longrightarrow R_3 : \quad \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0\\ 2 & 4 & 5 & 0 & 1 & 0\\ 0 & -1 & -3 & -3 & 0 & 1 \end{array}\right]$$

$$R_2 - 2R_1 \longrightarrow R_2 : \quad \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0\\ 0 & 0 & -1 & -2 & 1 & 0\\ 0 & -1 & -3 & -3 & 0 & 1 \end{array}\right]$$

$$\text{interchange } R_2 \text{ and } R_3 : \quad \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0\\ 0 & -1 & -3 & -3 & 0 & 1\\ 0 & 0 & -1 & -2 & 1 & 0 \end{array}\right]$$

$$R_2 \cdot (-1), \quad R_3 \cdot (-1): \qquad \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 3 & 3 & 0 & -1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right]$$

$$R_2 - 3R_3 \longrightarrow R_2: \qquad \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & -3 & 3 & -1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right]$$

$$R_1 - 2R_2 \longrightarrow R_1: \qquad \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 7 & -6 & 2 \\ 0 & 1 & 0 & -3 & 3 & -1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right]$$

$$R_1 - 3R_3 \longrightarrow R_1: \qquad \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -3 & 2 \\ 0 & 1 & 0 & -3 & 3 & -1 \\ 0 & 0 & 1 & 2 & -1 & 0 \end{array}\right]$$

So

$$\mathbf{A}^{-\mathbf{1}} = \begin{bmatrix} 1 & -3 & 2 \\ -3 & 3 & -1 \\ 2 & -1 & 0 \end{bmatrix}$$

Example 2. Find the inverse of

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ -2 & 1 & 8 \\ 1 & -2 & -7 \end{bmatrix}$$

or show that it does not exist.

Solution:

$$\text{form the augmented matrix } [\,\mathbf{A} \mid \mathbf{I}\,]: \qquad \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ -2 & 1 & 8 & 0 & 1 & 0 \\ 1 & -2 & -7 & 0 & 0 & 1 \end{array}\right]$$

$$R_3 - R_1 \longrightarrow R_3: \qquad \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ -2 & 1 & 8 & 0 & 1 & 0 \\ 0 & -4 & -8 & -1 & 0 & 1 \end{array}\right]$$

$$R_2 + 2R_1 \longrightarrow R_2: \qquad \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 5 & 10 & 2 & 1 & 0 \\ 0 & -4 & -8 & -1 & 0 & 1 \end{array}\right]$$

$$R_2/5, \quad R_3/(-4): \qquad \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2/5 & 1/5 & 0 \\ 0 & 1 & 2 & 1/4 & 0 & -1/4 \end{array}\right]$$

$$R_3 - R_2 \longrightarrow R_3 : \qquad \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2/5 & 1/5 & 0 \\ 0 & 0 & 0 & -3/20 & -1/5 & -1/4 \end{array}\right]$$

The row of zeroes on the left means we cannot get the identity matrix there, and thus $\mathbf{A}$ is singular (no inverse exists).

Applying this procedure to an arbitrary 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

we obtain (check!)

$$\mathbf{A}^{-1} = \frac{1}{\det\mathbf{A}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix},$$

where

$$\det\mathbf{A} = ad - bc,$$

provided that $\det\mathbf{A} \neq 0$. Otherwise, the inverse does not exist. In general, it is true that

$$\mathbf{A} \textit{ is invertible if and only if } \det\mathbf{A} \neq 0.$$

You can check that in the Example 2 above $\det\mathbf{A} = 0$.

3 Calculating determinants using row reduction

We can also use row reduction to compute large determinants. The idea is to use elementary row operations to reduce the matrix to an upper (or lower) triangular matrix, using the fact that

Determinant of an upper (lower) triangular or diagonal matrix equals the product of its diagonal entries.

As we row reduce, we need to keep in mind the following **properties of the determinants**:

1. $\det\mathbf{A} = \det\mathbf{A}^{\mathbf{T}}$, so we can apply either row or column operations to get the determinant.
2. If two rows or two columns of $\mathbf{A}$ are identical or if $\mathbf{A}$ has a row or a column of zeroes, then $\det\mathbf{A} = 0$.
3. If the matrix $\mathbf{B}$ is obtained by multiplying a *single* row or a single column of $\mathbf{A}$ by a number α, then
$$\det\mathbf{B} = \alpha\det\mathbf{A}.$$
   If *all* n rows (or all columns) of $\mathbf{A}$ are multiplied by α to obtain $\mathbf{B}$, then
$$\det\mathbf{B} = \alpha^n\det\mathbf{A}.$$

4. If $\mathbf{B}$ is obtained by interchanging two rows of $\mathbf{A}$, then
$$\det\mathbf{B} = -\det\mathbf{A}.$$

5. If $\mathbf{B}$ is obtained by adding a multiple of one row (or column) of $\mathbf{A}$ to another, then
$$\det\mathbf{B} = \det\mathbf{A}.$$

Example: use row reduction to compute the determinant of

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 3 & 1 \\ 0 & 4 & 3 & -3 \\ 2 & -1 & -1 & -3 \\ 0 & -4 & -3 & 2 \end{bmatrix} =$$

Solution:

Interchange R_2 and R_3: $$\det\mathbf{A} = -\begin{vmatrix} 2 & 3 & 3 & 1 \\ 2 & -1 & -1 & -3 \\ 0 & 4 & 3 & -3 \\ 0 & -4 & -3 & 2 \end{vmatrix} =$$

(note that the determinant changes sign, by property 4 above)

$$R_2 - R_1 \longrightarrow R_2 \qquad = -\begin{vmatrix} 2 & 3 & 3 & 1 \\ 0 & -4 & -4 & -4 \\ 0 & 4 & 3 & -3 \\ 0 & -4 & -3 & 2 \end{vmatrix} =$$

(determinant does not change)

$$R_4 + R_3 \longrightarrow R_4 \qquad = -\begin{vmatrix} 2 & 3 & 3 & 1 \\ 0 & -4 & -4 & -4 \\ 0 & 4 & 3 & -3 \\ 0 & 0 & 0 & -1 \end{vmatrix} =$$

(determinant does not change)

$$R_3 + R_2 \longrightarrow R_3 \qquad = -\begin{vmatrix} 2 & 3 & 3 & 1 \\ 0 & -4 & -4 & -4 \\ 0 & 0 & -1 & -7 \\ 0 & 0 & 0 & -1 \end{vmatrix} =$$

(determinant does not change, and we get an upper triangular matrix)

Compute the determinant of the upper triangular matrix: $$= -2 \cdot (-4) \cdot (-1) \cdot (-1) = 8$$

4 Homework problems

1. For each of the following matrices, find the inverse or show that it does not exist. In the latter case, check by calculating the determinant.

$$\text{a)}\begin{bmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \\ 1 & 1 & 2 \end{bmatrix} \qquad \text{b)}\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

$$\text{c)}\begin{bmatrix} 1 & -1 & -1 \\ 2 & 1 & 0 \\ 3 & -2 & 1 \end{bmatrix} \qquad \text{d)}\begin{bmatrix} 2 & 3 & 1 \\ -1 & 2 & 1 \\ 4 & -1 & -1 \end{bmatrix}$$

2. Use the method of row reduction to evaluate the following determinants:

$$\text{a)}\begin{vmatrix} 1 & 4 & 4 & 1 \\ 0 & 1 & -2 & 2 \\ 3 & 3 & 1 & 4 \\ 0 & 1 & -3 & -2 \end{vmatrix} \qquad \text{b)}\begin{vmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & -2 & 0 \\ -2 & 3 & -2 & 3 \\ 0 & -3 & 3 & 3 \end{vmatrix}$$

System of Linear Equations

A linear equation in variables $\boldsymbol{x_1, x_2, \dots, x_n}$ is an equation of the form

$$\boldsymbol{a_1x_1, a_2x_2, \dots, a_nx_n = b}$$

where $\boldsymbol{a_1, a_2, \dots, a_n}$ and $\boldsymbol{b}$ are constant real or complex numbers. The constant $\boldsymbol{a_i}$ is called the ***coefficient*** of $\boldsymbol{x_i}$; and $\boldsymbol{b}$ is called the ***constant term*** of the equation.

A system of linear equations (or linear system) is a finite collection of linear equations in samevariables. For instance, a linear system of $\boldsymbol{m}$ equations in $\boldsymbol{n}$ variables $\boldsymbol{x_1, x_2, \dots, x_n}$ can be written as

$$\begin{matrix} a_{11}x_1 & a_{12}x_2 & \cdots & a_{1n}x_n = b_1 \\ a_{21}x_1 & a_{22}x_2 & \cdots & a_{2n}x_n = b_2 \\ \vdots & \vdots & & \vdots \\ a_{m1}x_1 & a_{m2}x_2 & \cdots & a_{mn}x_n = b_m \end{matrix} \qquad \dots\dots\dots\dots\dots..(*)$$

A solution of a linear system (*) is a tuple $(\boldsymbol{s, s_2, \dots, s_n})$ of numbers that makes each equation a true statement when the values $\boldsymbol{s, s_2, \dots, s_n}$are substituted for $\boldsymbol{x_1, x_2, \dots, x_n}$, respectively. The set of all solutions of a linear system is called the solution set of the system.

Theorem 1.1. *Any system of linear equations has one of the following exclusive conclusions.*
(a) *No solution.*
(b) *Unique solution.*
(c) *Infinitely many solutions.*

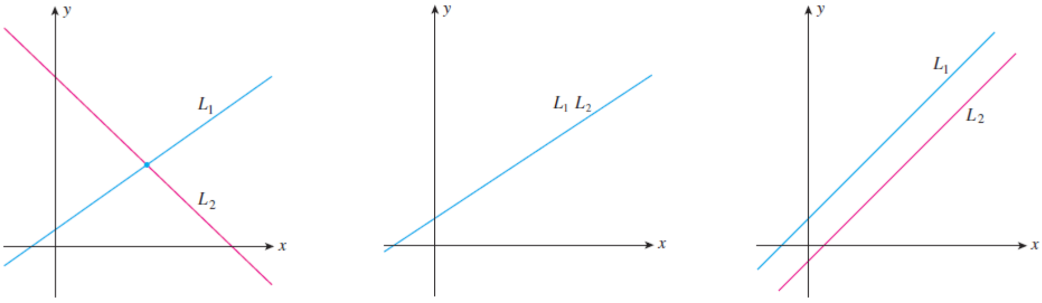


FIGURE 1
(a) Unique solution
(b) Infinitely many solutions
(c) No solution

A linear system is said to be **consistent** if it has at least one solution; and is said to be **inconsistent** if it has no solution.

We can represent the general system (*) by using matrices as $AX = B$ where $\boldsymbol{A}$ is called the ***coefficients matrix***, $\boldsymbol{X}$ is called the ***variables vector*** and $\boldsymbol{B}$ is called the ***constant vector***. Therefore the general system (*) can be rewrite as follows

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

$$A=\begin{bmatrix} a_{11} & a_{12} \cdots & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} \cdots & a_{mn} \end{bmatrix} ; X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \; and \; B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

The ***augmented*** matrix of the general linear system (*) is the matrix

$$[A:B] = \begin{bmatrix} a_{11} & a_{12} \cdots & a_{1n} & : & b_1 \\ a_{21} & a_{22} & a_{2n} & : & b_2 \\ \vdots & \ddots & \vdots & & \vdots \\ a_{m1} & a_{m2} \cdots & a_{mn} & : & b_m \end{bmatrix}$$

System of linear equations and matrices

It is impractical to solve more complicated linear systems by hand. Computers and calculators now have built in routines to solve larger and more complex systems. Matrices, in conjunction with graphing utilities and or computers are used for solving more complex systems. In this section, we will develop certain matrix methods for solving systems of linear equations.

There are many different methods to solve system of linear equations using matrices, in this section we will discuss some of these methods.

1) Cramer's Rule
This method used to solve square systems (number of equations equals the number of variables) *AX=B* and depends on the determinants.
   - If ***det(A)=0***, then the system has no solution.
   - If $\boldsymbol{det(A) \neq 0}$, then the system has a unique solution.

We will explain the method by the following examples

- Ex1. To use Cramer's Rule to solve the system

$$\begin{aligned} 3x_1 - 2x_2 &= 6 \\ -5x_1 + 4x_2 &= 8, \end{aligned}$$

we let

$$A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}, A_1 = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix}, \text{ and } A_2 = \begin{bmatrix} 3 & 6 \\ -5 & 8 \end{bmatrix}$$

where A_1 is matrix from A by replacing the first column by $\mathbf{b}$ and A_2 is matrix from A by replacing the second column by $\mathbf{b}$.

By Cramer's Rule, we obtain

$$x_1 = \frac{\det(A_1)}{\det(A)} = 20, \text{ and } x_2 = \frac{\det(A_2)}{\det(A)} = 27.$$

We can try another example with a 3 × 3 matrix:

- Ex2. To use Cramer's Rule to solve the system

$$\begin{aligned} x_1 - 2x_2 + \ x_3 &= \ 0 \\ 2x_2 - 8x_3 &= \ 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

we let

$$B = \begin{bmatrix} 1 & -2 & 1 \\ 0 & 2 & -8 \\ -4 & 5 & 9 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 8 \\ -9 \end{bmatrix}, B_1 = \begin{bmatrix} 0 & -2 & 1 \\ 8 & 2 & -8 \\ -9 & 5 & 9 \end{bmatrix},$$

$$B_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 8 & -8 \\ -4 & -9 & 9 \end{bmatrix}, \text{ and } B_3 = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 2 & 8 \\ -4 & 5 & -9 \end{bmatrix}$$

where B_i is matrix from B by replacing the column i by $\mathbf{b}$.

By Cramer's Rule, we obtain

$$x_1 = \frac{\det(B_1)}{\det(B)} = 29, x_2 = \frac{\det(B_2)}{\det(B)} = 16, \text{ and } x_3 = \frac{\det(B_3)}{\det(B)} = 3.$$

Besides solving system equation, we can use Cramer's rule to find the inverse matrix of a given matrix. Here we see an example:

2) **Elementary raw operations**

In this method we will used the elementary raw operations to change the augmented matrix $[A\colon B]$ to $[I\colon X]$. The following example explains this method.

Ex: Solve the following system of linear equation using the elementary row operations.

$$\begin{gathered} x-y+z=-4 \\ 2x-3y+4z=-15 \\ 5x+y-2z=12 \end{gathered}$$

Solution:

$$[A\colon B]=\begin{array}{ccc|c} 1 & -1 & 1 & -4 \\ 2 & -3 & 4 & -15 \\ 5 & 1 & -2 & 12 \end{array} \quad R_2\rightarrow -2r_1+r_2 \quad \begin{array}{ccc|c} 1 & -1 & 1 & -4 \\ 0 & -1 & 2 & -7 \\ 5 & 1 & -2 & 12 \end{array}$$

$$R_3\rightarrow -5r_1+r_3 \quad \begin{array}{ccc|c} 1 & -1 & 1 & -4 \\ 0 & -1 & 2 & -7 \\ 0 & 6 & -7 & 32 \end{array} \quad R_2\rightarrow -r_2 \quad \begin{array}{ccc|c} 1 & -1 & 1 & -4 \\ 0 & 1 & -2 & 7 \\ 0 & 6 & -7 & 32 \end{array}$$

$$R_3\rightarrow -6r_2+r_3 \quad \begin{array}{ccc|c} 1 & -1 & 1 & -4 \\ 0 & 1 & -2 & 7 \\ 0 & 0 & 5 & -10 \end{array} \quad R_3\rightarrow \tfrac{1}{5}r_3 \quad \begin{array}{ccc|c} 1 & -1 & 1 & -4 \\ 0 & 1 & -2 & 7 \\ 0 & 0 & 1 & -2 \end{array}$$

$$R_1\rightarrow r_2+r_1 \quad \begin{array}{ccc|c} 1 & 0 & -1 & 3 \\ 0 & 1 & -2 & 7 \\ 0 & 0 & 1 & -2 \end{array} \quad R_1\rightarrow r_3+r_1 \quad \begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & -2 & 7 \\ 0 & 0 & 1 & -2 \end{array}$$

$$R_2\rightarrow 2r_3+r_2 \quad \begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{array} \quad \textbf{Then the solution is} \begin{bmatrix} x=1 \\ y=3 \\ z=-2 \end{bmatrix}$$

Homework: Solve the following systems of linear equation using the elementary row operations.

$$\begin{gathered} 3x+y+4z=7 \\ 2x-3y+z=-8 \\ x+y-2z=1 \end{gathered}$$

$$\begin{gathered} x+2y-z=3 \\ 2x-y+2z=6 \\ x-3y+3z=4 \end{gathered}$$

Gaussian Elimination 30.2

Introduction

In this Section we will reconsider the Gaussian elimination approach discussed in HELM 8, and we will see how rounding error can grow if we are not careful in our implementation of the approach. A method called partial pivoting, which helps stop rounding error from growing, will be introduced.

Prerequisites

Before starting this Section you should . . .

- revise matrices, especially matrix solution of equations
- recall Gaussian elimination
- be able to find the inverse of a 2×2 matrix

Learning Outcomes

On completion you should be able to . . .

- carry out Gaussian elimination with partial pivoting

1. Gaussian elimination

Recall from HELM 8 that the basic idea with Gaussian (or Gauss) elimination is to replace the matrix of coefficients with a matrix that is easier to deal with. Usually the nicer matrix is of **upper triangular** form which allows us to find the solution by **back substitution**. For example, suppose we have

$$\begin{aligned} x_1+3x_2-5x_3 &= 2\\ 3x_1+11x_2-9x_3 &= 4\\ -x_1+x_2+6x_3 &= 5 \end{aligned}$$

which we can abbreviate using an **augmented matrix** to

$$\left[\begin{array}{rrr|r} \boxed{1} & 3 & -5 & 2\\ 3 & 11 & -9 & 4\\ -1 & 1 & 6 & 5 \end{array}\right].$$

We use the boxed element to eliminate any non-zeros below it. This involves the following row operations

$$\left[\begin{array}{rrr|r} \boxed{1} & 3 & -5 & 2\\ 3 & 11 & -9 & 4\\ -1 & 1 & 6 & 5 \end{array}\right] \begin{array}{l} \\ R2-3\times R1\\ R3+R1 \end{array} \Rightarrow \left[\begin{array}{rrr|r} \boxed{1} & 3 & -5 & 2\\ 0 & 2 & 6 & -2\\ 0 & 4 & 1 & 7 \end{array}\right].$$

And the next step is to use the $\boxed{2}$ to eliminate the non-zero below *it*. This requires the final row operation

$$\left[\begin{array}{rrr|r} 1 & 3 & -5 & 2\\ 0 & \boxed{2} & 6 & -2\\ 0 & 4 & 1 & 7 \end{array}\right] \begin{array}{l} \\ \\ R3-2\times R2 \end{array} \Rightarrow \left[\begin{array}{rrr|r} 1 & 3 & -5 & 2\\ 0 & \boxed{2} & 6 & -2\\ 0 & 0 & -11 & 11 \end{array}\right].$$

This is the augmented form for an upper triangular system, writing the system in extended form we have

$$\begin{aligned} x_1+3x_2-5x_3 &= 2\\ 2x_2+6x_3 &= -2\\ -11x_3 &= 11 \end{aligned}$$

which is easy to solve from the bottom up, by **back substitution**.

Example 5

Solve the system

$$\begin{aligned} x_1 + 3x_2 - 5x_3 &= 2 \\ 2x_2 + 6x_3 &= -2 \\ -11x_3 &= 11 \end{aligned}$$

Solution

The bottom equation implies that $x_3 = -1$. The middle equation then gives us that

$$2x_2 = -2 - 6x_3 = -2 + 6 = 4 \qquad \therefore \quad x_2 = 2$$

and finally, from the top equation,

$$x_1 = 2 - 3x_2 + 5x_3 = 2 - 6 - 5 = -9.$$

Therefore the solution to the problem stated at the beginning of this Section is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -9 \\ 2 \\ -1 \end{bmatrix}.$$

The following Task will act as useful revision of the Gaussian elimination procedure.

Task

Carry out row operations to reduce the matrix

$$\begin{bmatrix} 2 & -1 & 4 \\ 4 & 3 & -1 \\ -6 & 8 & -2 \end{bmatrix}$$

into upper triangular form.

Your solution

Answer

The row operations required to eliminate the non-zeros below the diagonal in the first column are as follows

$$\begin{bmatrix} 2 & -1 & 4 \\ 4 & 3 & -1 \\ -6 & 8 & -2 \end{bmatrix} \begin{matrix} \\ R2 - 2 \times R1 \\ R3 + 3 \times R1 \end{matrix} \Rightarrow \begin{bmatrix} 2 & -1 & 4 \\ 0 & 5 & -9 \\ 0 & 5 & 10 \end{bmatrix}$$

Next we use the 5 on the diagonal to eliminate the 5 below it:

$$\begin{bmatrix} 2 & -1 & 4 \\ 0 & 5 & -9 \\ 0 & 5 & 10 \end{bmatrix} \begin{matrix} \\ \\ R3 - R2 \end{matrix} \Rightarrow \begin{bmatrix} 2 & -1 & 4 \\ 0 & 5 & -9 \\ 0 & 0 & 19 \end{bmatrix}$$

which is in the required upper triangular form.

2. Partial pivoting

Partial pivoting is a refinement of the Gaussian elimination procedure which helps to prevent the growth of rounding error.

An example to motivate the idea

Consider the example

$$\begin{bmatrix} 10^{-4} & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

First of all let us work out the *exact* answer to this problem

$$\begin{aligned} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \begin{bmatrix} 10^{-4} & 1 \\ -1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{2 \times 10^{-4} + 1} \begin{bmatrix} 2 & -1 \\ 1 & 10^{-4} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ &= \frac{1}{2 \times 10^{-4} + 1} \begin{bmatrix} 1 \\ 1 + 10^{-4} \end{bmatrix} = \begin{bmatrix} 0.999800... \\ 0.999900... \end{bmatrix}. \end{aligned}$$

Now we compare this exact result with the output from Gaussian elimination. Let us suppose, for sake of argument, that all numbers are rounded to 3 significant figures. Eliminating the one non-zero element below the diagonal, and remembering that we are only dealing with 3 significant figures, we obtain

$$\begin{bmatrix} 10^{-4} & 1 \\ 0 & 10^4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 10^4 \end{bmatrix}.$$

The bottom equation gives $x_2 = 1$, and the top equation therefore gives $x_1 = 0$. Something has gone seriously wrong, for this value for x_1 is nowhere near the true value 0.9998…found without rounding.The problem has been caused by using a small number (10^{-4}) to eliminate a number much larger in magnitude (-1) below it.

The general idea with partial pivoting is to try to avoid using a small number to eliminate much larger numbers.

Suppose we swap the rows

$$\begin{bmatrix} -1 & 2 \\ 10^{-4} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and proceed as normal, still using just 3 significant figures. This time eliminating the non-zero below the diagonal gives

$$\begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

which leads to $x_2 = 1$ and $x_1 = 1$, which is an excellent approximation to the exact values, given that we are only using 3 significant figures.

Partial pivoting in general

At each step the aim in Gaussian elimination is to use an element on the diagonal to eliminate all the non-zeros below. In partial pivoting we look at all of these elements (the diagonal and the ones below) and swap the rows (if necessary) so that the element on the diagonal is not very much smaller than the other elements.

Key Point 3

Partial Pivoting

This involves scanning a column from the diagonal down. If the diagonal entry is very much smaller than any of the others we swap rows. Then we proceed with Gaussian elimination in the usual way.

In practice on a computer we swap rows to ensure that the diagonal entry is always the largest possible (in magnitude). For calculations we can carry out by hand it is usually only necessary to worry about partial pivoting if a zero crops up in a place which stops Gaussian elimination working. Consider this example

$$\begin{bmatrix} 1 & -3 & 2 & 1 \\ 2 & -6 & 1 & 4 \\ -1 & 2 & 3 & 4 \\ 0 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -4 \\ 1 \\ 12 \\ 0 \end{bmatrix}.$$

The first step is to use the 1 in the top left corner to eliminate all the non-zeros below it in the augmented matrix

$$\left[\begin{array}{cccc|c} 1 & -3 & 2 & 1 & -4 \\ 2 & -6 & 1 & 4 & 1 \\ -1 & 2 & 3 & 4 & 12 \\ 0 & -1 & 1 & 1 & 0 \end{array}\right] \begin{array}{l} \\ R2 - 2 \times R1 \\ R3 + R1 \\ \\ \end{array} \Rightarrow \left[\begin{array}{cccc|c} 1 & -3 & 2 & 1 & -4 \\ 0 & \boxed{0} & -3 & 2 & 9 \\ 0 & -1 & 5 & 5 & 8 \\ 0 & -1 & 1 & 1 & 0 \end{array}\right].$$

What we would *like* to do now is to use the boxed element to eliminate all the non-zeros below it. But clearly this is impossible. We need to apply partial pivoting. We look **down** the column starting at the diagonal entry and see that the two possible candidates for the swap are both equal to -1. Either will do so let us swap the second and fourth rows to give

$$\left[\begin{array}{cccc|c} 1 & -3 & 2 & 1 & -4 \\ 0 & \boxed{-1} & 1 & 1 & 0 \\ 0 & -1 & 5 & 5 & 8 \\ 0 & 0 & -3 & 2 & 9 \end{array}\right].$$

That was the partial pivoting step. Now we proceed with Gaussian elimination

$$\left[\begin{array}{cccc|c} 1 & -3 & 2 & 1 & -4 \\ 0 & \boxed{-1} & 1 & 1 & 0 \\ 0 & -1 & 5 & 5 & 8 \\ 0 & 0 & -3 & 2 & 9 \end{array}\right] \begin{array}{c} \\ \\ R3 - R2 \\ \\ \end{array} \Rightarrow \left[\begin{array}{cccc|c} 1 & -3 & 2 & 1 & -4 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 4 & 4 & 8 \\ 0 & 0 & -3 & 2 & 9 \end{array}\right].$$

The arithmetic is simpler if we cancel a factor of 4 out of the third row to give

$$\left[\begin{array}{cccc|c} 1 & -3 & 2 & 1 & -4 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & -3 & 2 & 9 \end{array}\right].$$

And the elimination phase is completed by removing the -3 from the final row as follows

$$\left[\begin{array}{cccc|c} 1 & -3 & 2 & 1 & -4 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & \boxed{1} & 1 & 2 \\ 0 & 0 & -3 & 2 & 9 \end{array}\right] \begin{array}{c} \\ \\ \\ R4 + 3 \times R3 \end{array} \Rightarrow \left[\begin{array}{cccc|c} 1 & -3 & 2 & 1 & -4 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 5 & 15 \end{array}\right].$$

This system is upper triangular so back substitution can be used now to work out that $x_4 = 3$, $x_3 = -1$, $x_2 = 2$ and $x_1 = 1$.

The Task below is a case in which partial pivoting is required.

[For a large system which can be solved by Gauss elimination see Engineering Example 1 on page 62].

Task

Transform the matrix

$$\left[\begin{array}{ccc} 1 & -2 & 4 \\ -3 & 6 & -11 \\ 4 & 3 & 5 \end{array}\right]$$

into upper triangular form using Gaussian elimination (with partial pivoting when necessary).

Your solution

Answer

The row operations required to eliminate the non-zeros below the diagonal in the first column are

$$\begin{bmatrix} 1 & -2 & 4 \\ -3 & 6 & -11 \\ 4 & 3 & 5 \end{bmatrix} \begin{matrix} \\ R2+3\times R1 \\ R3-4\times R1 \end{matrix} \Rightarrow \begin{bmatrix} 1 & -2 & 4 \\ 0 & 0 & 1 \\ 0 & 11 & -11 \end{bmatrix}$$

which puts a zero on the diagonal. We are forced to use partial pivoting and swapping the second and third rows gives

$$\begin{bmatrix} 1 & -2 & 4 \\ 0 & 11 & -11 \\ 0 & 0 & 1 \end{bmatrix}$$

which is in the required upper triangular form.

Key Point 4

When To Use Partial Pivoting

1. When carrying out Gaussian elimination on a computer, we would usually always swap rows so that the element on the diagonal is as large (in magnitude) as possible. This helps stop the growth of rounding error.
2. When doing hand calculations (not involving rounding) there are two reasons we might pivot
   - (a) If the element on the diagonal is zero, we **have** to swap rows so as to put a non-zero on the diagonal.
   - (b) Sometimes we might swap rows so that there is a "nicer" non-zero number on the diagonal than there would be without pivoting. For example, if the number on the diagonal can be arranged to be a 1 then no awkward fractions will be introduced when we carry out row operations related to Gaussian elimination.

Exercises

1. Solve the following system by back substitution

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 3 \\ 5x_2 + 6x_3 &= -2 \\ 7x_3 &= -14 \end{aligned}$$

2. (a) Show that the exact solution of the system of equations

$$\begin{bmatrix} 10^{-5} & 1 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 10 \end{bmatrix} \text{ is } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -0.99998 \\ 2.00001 \end{bmatrix}.$$

(b) Working to 3 significant figures, and using Gaussian elimination *without* pivoting, find an approximation to $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Show that the rounding error causes the approximation to x_1 to be a very poor one.

(c) Working to 3 significant figures, and using Gaussian elimination *with* pivoting, find an approximation to $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Show that the approximation this time is a good one.

3. Carry out row operations (with partial pivoting if necessary) to reduce these matrices to upper triangular form.

$$\text{(a)} \begin{bmatrix} 1 & -2 & 4 \\ -4 & -3 & -3 \\ -1 & 13 & 1 \end{bmatrix}, \quad \text{(b)} \begin{bmatrix} 0 & -1 & 2 \\ 1 & -4 & 2 \\ -2 & 5 & -4 \end{bmatrix}, \quad \text{(c)} \begin{bmatrix} -3 & 10 & 1 \\ 1 & -3 & 2 \\ -2 & 10 & -4 \end{bmatrix}.$$

(Hint: before tackling (c) you might like to consider point 2(b) in Key Point 4.)

Answers

1. From the last equation we see that $x_3 = -2$. Using this information in the second equation gives us $x_2 = 2$. Finally, the first equation implies that $x_1 = -3$.

2. (a) The formula $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \dfrac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ can be used to show that

$$x_1 = -\frac{50000}{50001} = -0.99998 \qquad \text{and} \qquad x_2 = \frac{200005}{100002} = 2.00001 \qquad \text{as required.}$$

(b) Carrying out the elimination without pivoting, and rounding to 3 significant figures we find that $x_2 = 2.00$ and that, therefore, $x_1 = 0$. This is a very poor approximation to x_1.

(c) To apply partial pivoting we swap the two rows and then eliminate the bottom left element. Consequently we find that, after rounding the system of equations to 3 significant figures, $x_2 = 2.00$ and $x_1 = -1.00$. These give excellent agreement with the exact answers.

Answers

3.

(a) The row operations required to eliminate the non-zeros below the diagonal in the first column are as follows

$$\begin{bmatrix} 1 & -2 & 4 \\ -4 & -3 & -3 \\ -1 & 13 & 1 \end{bmatrix} \begin{matrix} \\ R2+4\times R1 \\ R3+1\times R1 \end{matrix} \Rightarrow \begin{bmatrix} 1 & -2 & 4 \\ 0 & -11 & 13 \\ 0 & 11 & 5 \end{bmatrix}$$

Next we use the element in the middle of the matrix to eliminate the value underneath it. This gives

$$\begin{bmatrix} 1 & -2 & 4 \\ 0 & -11 & 13 \\ 0 & 0 & 18 \end{bmatrix} \qquad \text{which is of the required upper triangular form.}$$

(b) We must swap the rows to put a non-zero in the top left position (this is the partial pivoting step). Swapping the first and second rows gives the matrix

$$\begin{bmatrix} 1 & -4 & 2 \\ 0 & -1 & 2 \\ -2 & 5 & -4 \end{bmatrix}.$$

We carry out one row operation to eliminate the non-zero in the bottom left entry as follows

$$\begin{bmatrix} 1 & -4 & 2 \\ 0 & -1 & 2 \\ -2 & 5 & -4 \end{bmatrix} \begin{matrix} \\ \\ R3+2\times R1 \end{matrix} \Rightarrow \begin{bmatrix} 1 & -4 & 2 \\ 0 & -1 & 2 \\ 0 & -3 & 0 \end{bmatrix}$$

Next we use the middle element to eliminate the non-zero value underneath it. This gives

$$\begin{bmatrix} 1 & -4 & 2 \\ 0 & -1 & 2 \\ 0 & 0 & -6 \end{bmatrix} \qquad \text{which is of the required upper triangular form.}$$

(c) If we swap the first and second rows of the matrix then we do not have to deal with fractions. Having done this the row operations required to eliminate the non-zeros below the diagonal in the first column are as follows

$$\begin{bmatrix} 1 & -3 & 2 \\ -3 & 10 & 1 \\ -2 & 10 & -4 \end{bmatrix} \begin{matrix} \\ R2+3\times R1 \\ R3+2\times R1 \end{matrix} \Rightarrow \begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & 7 \\ 0 & 4 & 0 \end{bmatrix}$$

Next we use the element in the middle of the matrix to eliminate the non-zero value underneath it. This gives

$$\begin{bmatrix} 1 & -3 & 2 \\ 0 & 1 & 7 \\ 0 & 0 & -28 \end{bmatrix} \qquad \text{which is of the required upper triangular form.}$$

Gauss-Jordan Matrix Elimination

-This method can be used to solve systems of **linear equations** involving two or more variables. However, the system must be changed to an augmented matrix.

-This method can also be used to find the inverse of a 2x2 matrix or larger matrices, 3x3, 4x4 etc.

Note: The matrix must be a square matrix in order to find its inverse.

An **Augmented Matrix** is used to solve a system of linear equations.

$$\text{System of Equations} \longrightarrow \quad \begin{aligned} a_1 x + b_1 y + c_1 z &= d_1 \\ a_2 x + b_2 y + c_2 z &= d_2 \\ a_3 x + b_3 y + c_3 z &= d_3 \end{aligned}$$

$$\text{Augmented Matrix} \longrightarrow \quad \left[\begin{array}{ccc||c} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{array}\right]$$

-When given a system of equations, to write in augmented matrix form, the coefficients of each variable must be taken and put in a matrix.

For example, for the following system:

$$\begin{aligned} 3x + 2y - z &= 3 \\ x - y + 2z &= 4 \\ 2x + 3y - z &= 3 \end{aligned}$$

$$\text{Augmented Matrix} \longrightarrow \quad \left[\begin{array}{ccc||c} 3 & 2 & -1 & 3 \\ 1 & -1 & 2 & 4 \\ 2 & 3 & -1 & 3 \end{array}\right]$$

-There are three different operations known as **Elementary Row Operations** used when solving or reducing a matrix, using Gauss-Jordan elimination method.

1. Interchanging two rows.
2. Add one row to another row, or multiply one row first and then adding it to another.
3. Multiplying a row by any constant greater than zero.

Identity Matrix-is the final result obtained when a matrix is reduced. This matrix consists of ones in the diagonal starting with the first number.

-The numbers in the last column are the answers to the system of equations.

$$\left[\begin{array}{ccc||c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 5 \end{array}\right] \longleftarrow \text{Identity Matrix for a 3x3}$$

$$\left[\begin{array}{cccc||c} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 4 \end{array}\right] \longleftarrow \text{Identity Matrix for a 4x4}$$

-The pattern continues for bigger matrices.

Solving a system using Gauss-Jordan

–The best way to go is to get the ones first in their respective column, and then using that one to get the zeros in that column.

-It is very important to understand that there is no exact procedure to follow when using the Gauss-Jordan method to solve for a system.

$$\begin{aligned} 3x + 2y - z &= 3 \\ x - y + 2z &= 4 \\ 2x + 3y - z &= 3 \end{aligned}$$

Write as an augmented matrix.

$$\downarrow$$

$$\left[\begin{array}{ccc||c} 3 & 2 & -1 & 3 \\ 1 & -1 & 2 & 4 \\ 2 & 3 & -1 & 3 \end{array}\right]$$

Switch row 1 with row 2 to get a 1 in the first column

$$\downarrow$$

$$\left[\begin{array}{ccc||c} 1 & -1 & 2 & 4 \\ 3 & 2 & -1 & 3 \\ 2 & 3 & -1 & 3 \end{array}\right]$$ ***Multiply row 1 by -3 and add to row 2 to get a zero***

Row 1 multiplied by -3 ⟶		-3	3	-6	-12
Row 2 ⟶	$+$	3	2	-1	3
New Row 2 ⟶		0	5	-7	-9

-Put the new row 2 in the matrix, note that though row 1 was multiplied by -3, row 1 didn't change in our matrix.

$$\left[\begin{array}{ccc||c} 1 & -1 & 2 & 4 \\ 0 & 5 & -7 & -9 \\ 2 & 3 & -1 & 3 \end{array}\right]$$

Using a similar procedure of multiplying and adding rows, obtain the following matrix

$$\left[\begin{array}{ccc|c} 1 & -1 & 2 & 4 \\ 0 & 5 & -7 & -9 \\ 2 & 3 & -1 & 3 \end{array}\right]$$ ***Multiply row 1 by -2 and add to row3 as above.***

$$\downarrow$$

$$\left[\begin{array}{ccc||c} 1 & -1 & 2 & 4 \\ 0 & 5 & -7 & -9 \\ 0 & 5 & -5 & -5 \end{array}\right]$$ ***Switch rows 2 and 3 to obtain the following***

$$\downarrow$$

$$\left[\begin{array}{ccc||c} 1 & -1 & 2 & 4 \\ 0 & 5 & -5 & -5 \\ 0 & 5 & -7 & -9 \end{array}\right]$$ ***Divide the second row by 5 to obtain a 1 in the second row.***

$$\downarrow$$

$$\left[\begin{array}{ccc||c} 1 & -1 & 2 & 4 \\ 0 & 1 & -1 & -1 \\ 0 & 5 & -7 & -9 \end{array}\right]$$ ***Add row 2 to row 1***

$$\downarrow$$

$$\left[\begin{array}{ccc||c} 1 & 0 & 1 & 3 \\ 0 & 1 & -1 & -1 \\ 0 & 5 & -7 & -9 \end{array}\right]$$ ***Multiply and add like we did earlier*, *-5 * R2+R3***

$$\downarrow$$

$$\left[\begin{array}{ccc|c}1 & 0 & 1 & 3\\ 0 & 1 & -1 & -1\\ 0 & 0 & -2 & -4\end{array}\right]$$ ***Divide row 3 by -2 to obtain a 1 in the third row.***

$$\downarrow$$

$$\left[\begin{array}{ccc|c}1 & 0 & 1 & 3\\ 0 & 1 & -1 & -1\\ 0 & 0 & 1 & 2\end{array}\right]$$

-Finally, the matrix can be solved in two different ways:
A. Using the *1* in column 3, obtain the other zeros and the solutions.

$$\left[\begin{array}{ccc|c}1 & 0 & 0 & 1\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 2\end{array}\right] \quad x=1 \quad y=1 \quad z=2$$

B. Solve by using back substitution.

-The solution to the last row is $z = 2$, the answer can be substituted into the equation produced by the second row. $y - z = -1$ Substituting into this equation, it simplifies to:

$$y-2=-1$$
$$y=1$$

-Again, substituting the answer for z into the first equation will give the answer for x.

$$x+z=3$$
$$x+2=3$$
$$x=1$$

Vector Spaces

4.1 Vectors in R^n

- An ordered n-tuple :

  a sequence of n real numbers $(x_1, x_2, \cdots, x_n)$

- R^n-space :

  the set of all ordered n-tuples

$n = 1$ R^1-space = set of all real numbers

(R^1-space can be represented geometrically by the x-axis)

$n = 2$ R^2-space = set of all ordered pair of real numbers (x_1, x_2)

(R^2-space can be represented geometrically by the xy-plane)

$n = 3$ R^3-space = set of all ordered triple of real numbers (x_1, x_2, x_3)

(R^3-space can be represented geometrically by the xyz-space)

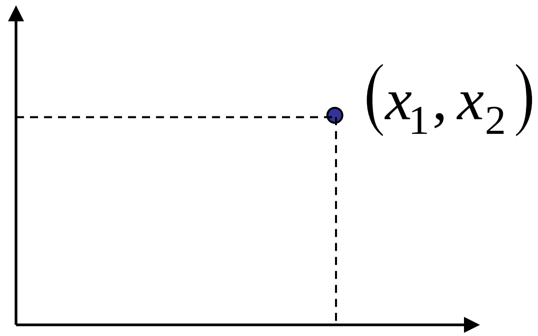
$n = 4$ R^4-space = set of all ordered quadruple of real numbers (x_1, x_2, x_3, x_4)

- Notes:

  (1) An n-tuple $(x_1, x_2, \cdots, x_n)$ can be viewed as a point in R^n with the x_i's as its coordinates

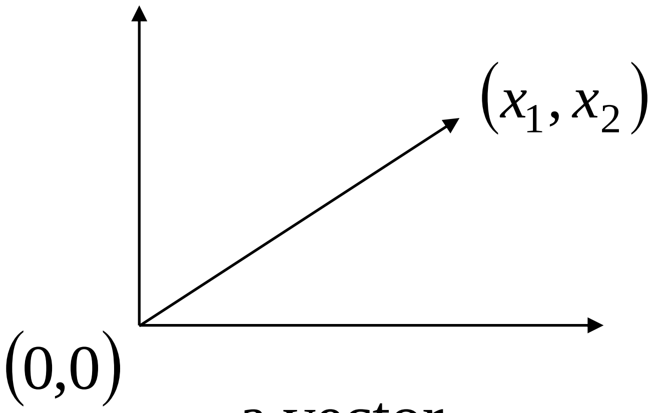
  (2) An n-tuple $(x_1, x_2, \cdots, x_n)$ also can be viewed as a vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ in R^n with the x_i's as its components

- Ex:



a point

or



a vector

※ A vector on the plane is expressed geometrically by a directed line segment whose initial point is the origin and whose terminal point is the point (x_1, x_2)

$$\mathbf{u}=(u_1,u_2,\cdots,u_n),\quad \mathbf{v}=(v_1,v_2,\cdots,v_n)$$ (two vectors in R^n)

- Equality:

$\mathbf{u}=\mathbf{v}$ if and only if $u_1=v_1,\ u_2=v_2,\ \cdots,\ u_n=v_n$

- Vector addition (the sum of **u** and **v**):

$$\mathbf{u}+\mathbf{v}=(u_1+v_1,u_2+v_2,\cdots,u_n+v_n)$$

- Scalar multiplication (the scalar multiple of **u** by c):

$$c\mathbf{u}=(cu_1,cu_2,\cdots,cu_n)$$

- Notes:

The sum of two vectors and the scalar multiple of a vector in R^n are called the standard operations in R^n

- Difference between **u** and **v**:

$$\mathbf{u}-\mathbf{v}\equiv\mathbf{u}+(-1)\mathbf{v}=(u_1-v_1,\ u_2-v_2,\ u_3-v_3,\ldots,\ u_n-v_n)$$

- Zero vector

$$\mathbf{0}=(0, 0, \ldots, 0)$$

- Notes:

A vector $\mathbf{u}=(u_1,u_2,\ldots,u_n)$ in R^n can be viewed as:

Use comma to separate components

a $1\times n$ row matrix (row vector): $\mathbf{u}=[u_1\ u_2\ \cdots\ u_n]$

Use blank space to separate entries

or

a $n\times 1$ column matrix (column vector): $\mathbf{u}=\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$

※ Therefore, the operations of matrix addition and scalar multiplication generate the same results as the corresponding vector operations (see the next slide)

Vector addition

$$\mathbf{u}+\mathbf{v}=(u_1,u_2,\cdots,u_n)+(v_1,v_2,\cdots,v_n)$$
$$=(u_1+v_1,u_2+v_2,\cdots,u_n+v_n)$$

Scalar multiplication

$$c\mathbf{u}=c(u_1,\ u_2,\cdots,\ u_n)$$
$$=(cu_1,cu_2,\cdots,\ cu_n)$$

Regarded as 1×*n* row matrix

$$\mathbf{u}+\mathbf{v}=[u_1\ \ u_2\ \ \cdots\ \ u_n]+[v_1\ \ v_2\ \ \cdots\ \ v_n]$$
$$=[u_1+v_1\ \ u_2+v_2\ \ \cdots\ \ u_n+v_n]$$

$$c\mathbf{u}=c[u_1\ u_2\cdots u_n]$$
$$=[cu_1\ \ cu_2\ \ \cdots\ \ cu_n]$$

Regarded as *n*×1 column matrix

$$\mathbf{u}+\mathbf{v}=\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix}+\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}=\begin{bmatrix}u_1+v_1\\u_2+v_2\\\vdots\\u_n+v_n\end{bmatrix}$$

$$c\mathbf{u}=c\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix}=\begin{bmatrix}cu_1\\cu_2\\\vdots\\cu_n\end{bmatrix}$$

- Properties of vector addition and scalar multiplication

Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in R^n, and let c and d be scalars

(1) $\mathbf{u}+\mathbf{v}$ is a vector in R^n (closure under vector addition)

(2) $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$ (commutative property of vector addition)

(3) $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$ (associative property of vector addition)

(4) $\mathbf{u}+\mathbf{0} = \mathbf{u}$ (additive identity property)

(5) $\mathbf{u}+(-\mathbf{u}) = \mathbf{0}$ (additive inverse property) (Note that $-\mathbf{u}$ is just the notation of the additive inverse of $\mathbf{u}$, and $-\mathbf{u} = (-1)\mathbf{u}$ will be proved next)

(6) $c\mathbf{u}$ is a vector in R^n (closure under scalar multiplication)

(7) $c(\mathbf{u}+\mathbf{v}) = c\mathbf{u}+c\mathbf{v}$ (distributive property of scalar multiplication over vector addition)

(8) $(c+d)\,\mathbf{u} = c\mathbf{u}+d\mathbf{u}$ (distributive property of scalar multiplication over real-number addition)

(9) $c\,(d\,\mathbf{u}) = (cd)\mathbf{u}$ (associative property of multiplication)

(10) $1(\mathbf{u}) = \mathbf{u}$ (multiplicative identity property)

- Ex 5: Practice standard vector operations in R^4

Let $\mathbf{u}$ = (2, –1, 5, 0), $\mathbf{v}$ = (4, 3, 1, –1), and $\mathbf{w}$ = (–6, 2, 0, 3) be vectors in R^4. Solve $\mathbf{x}$ in each of the following cases.

(a) $\mathbf{x} = 2\mathbf{u} - (\mathbf{v} + 3\mathbf{w})$

(b) $3(\mathbf{x}+\mathbf{w}) = 2\mathbf{u} - \mathbf{v}+\mathbf{x}$

Sol: (a)

$$\begin{aligned}\mathbf{x} &= 2\mathbf{u}-(\mathbf{v}+3\mathbf{w})\\ &= 2\mathbf{u}+(-1)(\mathbf{v}+3\mathbf{w})\\ &= 2\mathbf{u}-\mathbf{v}-3\mathbf{w}\\ &= (4,\ -2,\ 10,\ 0)-(4,\ 3,\ 1,\ -1)-(-18,\ 6,\ 0,\ 9)\\ &= (4-4+18,\ -2-3-6,\ 10-1-0,\ 0+1-9)\\ &= (18,\ -11,\ 9,\ -8)\end{aligned}$$

(b) $$\begin{aligned} 3(\mathbf{x}+\mathbf{w}) &= 2\mathbf{u}-\mathbf{v}+\mathbf{x} \\ 3\mathbf{x}+3\mathbf{w} &= 2\mathbf{u}-\mathbf{v}+\mathbf{x} \\ 3\mathbf{x}-\mathbf{x} &= 2\mathbf{u}-\mathbf{v}-3\mathbf{w} \\ 2\mathbf{x} &= 2\mathbf{u}-\mathbf{v}-3\mathbf{w} \\ \mathbf{x} &= \mathbf{u}-\tfrac{1}{2}\mathbf{v}-\tfrac{3}{2}\mathbf{w} \\ &= \left(2,-1,5,0\right)+\left(-2,\tfrac{-3}{2},\tfrac{-1}{2},\tfrac{1}{2}\right)+\left(9,-3,0,\tfrac{-9}{2}\right) \\ &= \left(9,\tfrac{-11}{2},\tfrac{9}{2},-4\right) \end{aligned}$$

- Notes:

  (1) The zero vector **0** in R^n is called the additive identity in R^n

  (2) The vector –**u** is called the additive inverse of **u**

- (Properties of additive identity and additive inverse)

  Let **v** be a vector in R^n and c be a scalar. Then the following properties are true

(1) The additive identity is unique, i.e., if **v**+**u** = **v**, **u** must be **0**

(2) The additive inverse of **v** is unique, i.e., if **v**+**u** = **0**, **u** must be –**v**

(3) 0**v** = **0**

(4) c**0** = **0**

(5) If c**v** = **0**, either c = 0 or **v** = **0**

(6) –(–**v**) = **v**

- Linear combination (線性組合) in R^n:

The vector **x** is called a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, if it can be expressed in the form

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n,\ \text{where } c_1,\ c_2, \ldots,\ c_n \text{ are real numbers}$$

- Ex 6:

Given **x** = (–1, –2, –2), **u** = (0,1,4), **v** = (– 1,1,2), and **w** = (3,1,2) in R^3, find *a*, *b*, and *c* such that **x** = *a***u** + *b***v** + *c***w**.

Sol:

$$\begin{array}{rcrcrcr} & & -b & + & 3c & = & -1 \\ a & + & b & + & c & = & -2 \\ 4a & + & 2b & + & 2c & = & -2 \end{array}$$

$$\Rightarrow a = 1,\ b = -2,\ c = -1$$

Thus $\mathbf{x} = \mathbf{u} - 2\mathbf{v} - \mathbf{w}$

4.2 Vector Spaces

- Vector spaces

Let V be a set on which two operations (addition and scalar multiplication) are defined. **If the following ten axioms are satisfied** for every element **u**, **v**, and **w** in V and every scalar (real number) c and d, then **V is called a vector space**, and the **elements** in V are called **vectors**

Addition:

(1) **u**+**v** is in V

(2) **u**+**v** = **v**+**u**

(3) **u**+(**v**+**w**) = (**u**+**v**)+**w**

(4) V has a zero vector **0** such that for every **u** in V, **u**+**0** = **u**

(5) For every **u** in V, there is a vector in V denoted by –**u** such that **u**+(–**u**) = **0**

Scalar multiplication:

(6) $c\mathbf{u}$ is in V

(7) $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$

(8) $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$

(9) $c(d\mathbf{u})=(cd)\mathbf{u}$

(10) $1(\mathbf{u})=\mathbf{u}$

- Notes:

A vector space consists of four entities:

a set of vectors, a set of real-number scalars, and two operations

V: nonempty set of vectors

c: any scalar

$+(\mathbf{u},\mathbf{v})=\mathbf{u}+\mathbf{v}$: vector addition

$\cdot(c,\mathbf{u})=c\mathbf{u}$: scalar multiplication

$(V, \ +, \ \cdot)$ is called a vector space

※ The set V together with the definitions of vector addition and scalar multiplication satisfying the above ten axioms is called a vector space

- Four examples of vector spaces are shown as follows.

(1) *n*-tuple space: R^n

$$(u_1,u_2,\cdots u_n)+(v_1,v_2,\cdots v_2)=(u_1+v_1,u_2+v_2,\cdots u_n+v_n)$$ (standard vector addition)

$$k(u_1,u_2,\cdots u_n)=(ku_1,ku_2,\cdots ku_n)$$ (standard scalar multiplication for vectors)

(2) Matrix space): $V = M_{m\times n}$

(the set of all $m\times n$ matrices with real-number entries)

Ex: $(m = n = 2)$

$$\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}+\begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}=\begin{bmatrix} u_{11}+v_{11} & u_{12}+v_{12} \\ u_{21}+v_{21} & u_{22}+v_{22} \end{bmatrix}$$ (standard matrix addition)

$$k\begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}=\begin{bmatrix} ku_{11} & ku_{12} \\ ku_{21} & ku_{22} \end{bmatrix}$$ (standard scalar multiplication for matrices)

(3) *n*-th degree or less polynomial space: $V = P_n$

(the set of all real-valued polynomials of degree *n* **or less**)

$$p(x)+q(x)=(a_0+b_0)+(a_1+b_1)x+\cdots+(a_n+b_n)x^n$$ (standard polynomial addition)

$$kp(x)=ka_0+ka_1x+\cdots+ka_nx^n$$ (standard scalar multiplication for polynomials)

※ By the fact that the set of real numbers is closed under addition and multiplication, it is straightforward to show that P_n satisfies the ten axioms and thus is a vector space

(4) Continuous function space : $V = C(-\infty,\infty)$

(the set of all real-valued continuous functions defined on the entire real line)

$$(f+g)(x)=f(x)+g(x)$$ (standard addition for functions)

$$(kf)(x)=kf(x)$$ (standard scalar multiplication for functions)

※ By the fact that the sum of two continuous function is continuous and the product of a scalar and a continuous function is still a continuous function, $C(-\infty,\infty)$ is a vector space

- Summary of important vector spaces

R = set of all real numbers

R^2 = set of all ordered pairs

R^3 = set of all ordered triples

R^n = set of all n-tuples

$C(-\infty, \infty)$ = set of all continuous functions defined on the real number line

$C[a,b]$ = set of all continuous functions defined on a closed interval $[a,b]$

P = set of all polynomials

P_n = set of all polynomials of degree $\leq n$

$M_{m,n}$ = set of $m \times n$ matrices

$M_{n,n}$ = set of $n \times n$ square matrices

- Notes: To show that a set is not a vector space, you need only find one axiom that is not satisfied

- Ex 6: The set of all integers is not a vector space

  Pf: $1 \in V$, and $\tfrac{1}{2}$ is a real-number scalar

$$\underset{\substack{\uparrow\\ \text{scalar}}}{(\tfrac{1}{2})}\underset{\substack{\uparrow\\ \text{integer}}}{(1)} = \underset{\substack{\uparrow\\ \text{noninteger}}}{\tfrac{1}{2}} \notin V$$

(it is not closed under scalar multiplication)

- Ex 7: The set of all (exact) second-degree polynomial functions is not a vector space

  Pf: Let $p(x) = x^2$ and $q(x) = -x^2 + x + 1$

$$\Rightarrow p(x) + q(x) = x + 1 \notin V$$

(it is not closed under vector addition)

- Ex 8:

$V = R^2 =$ the set of all ordered pairs of real numbers

vector addition : $(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$

scalar multiplication: $c(u_1, u_2) = (cu_1, 0)$ (nonstandard definition)

Verify V is not a vector space

Sol:

This kind of setting can satisfy the first nine axioms of the definition of a vector space (you can try to show that), but it violates the tenth axiom

$\because 1(1, 1) = (1, 0) \neq (1, 1)$

$\therefore$ the set (together with the two given operations) is not a vector space

4.3 Subspaces of Vector Spaces

- Subspace

$(V,+,\cdot)$: a vector space

$$\left.\begin{array}{l} W \neq \Phi \\ W \subseteq V \end{array}\right\} : \text{a nonempty subset of } V$$

$(W,+,\cdot)$: The nonempty subset W is called a subspace **if** W **is a vector space** under the operations of vector addition and scalar multiplication defined on V

- Trivial subspace

- Every vector space V has at least two subspaces

(1) Zero vector space $\{\mathbf{0}\}$ is a subspace of V (It satisfies the ten axioms)

(2) V is a subspace of V

※ Any subspaces other than these two are called proper (or nontrivial) subspaces

- Examination of whether W being a subspace
  - Since the vector operations defined on W are the same as those defined on V, and most of the ten axioms inherit the properties for the vector operations, it is not needed to verify those axioms
  - To identify that a nonempty subset of a vector space is a subspace, it is sufficient to **test only the closure conditions under vector addition and scalar multiplication**

- Test whether a nonempty subset being a subspace

If W is a nonempty subset of a vector space V, then W is a subspace of V if and only if the following conditions hold

(1) If $\mathbf{u}$ and $\mathbf{v}$ are in W, then $\mathbf{u}+\mathbf{v}$ is in W

(2) If $\mathbf{u}$ is in W and c is any scalar, then $c\mathbf{u}$ is in W

- Ex 2: A subspace of $M_{2\times2}$

Let W be the set of all 2×2 symmetric matrices. Show that W is a subspace of the vector space $M_{2\times2}$, with the standard operations of matrix addition and scalar multiplication

Sol:

First, we knon that W, the set of all 2×2 symmetric matrices, is an nonempty subset of the vector space $M_{2\times2}$

Second,

$$A_1 \in W, A_2 \in W \Rightarrow (A_1 + A_2)^T = A_1^T + A_2^T = A_1 + A_2 \quad (A_1 + A_2 \in W)$$

$$c \in R,\ A \in W \Rightarrow (cA)^T = cA^T = cA \quad (cA \in W)$$

The definition of a symmetric matrix A is that $A^T = A$

Thus, Thm. 4.5 is applied to obtain that W is a subspace of $M_{2\times2}$

- Ex 3: The set of singular matrices is not a subspace of $M_{2\times 2}$

Let W be the set of singular (noninvertible) matrices of order 2. Show that W is not a subspace of $M_{2\times 2}$ with the standard matrix operations

Sol:

$$A=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\in W,\ B=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\in W$$

$$\because A+B=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=I\notin W$$ (W is not closed under vector addition)

$$\therefore W \text{ is not a subspace of } M_{2\times 2}$$

- Ex 4: The set of first-quadrant vectors is not a subspace of R^2

Show that $W = \{(x_1, x_2) : x_1 \geq 0 \text{ and } x_2 \geq 0\}$, with the standard operations, is not a subspace of R^2

Sol:

Let $\mathbf{u} = (1,\ 1) \in W$

$$\because (-1)\mathbf{u} = (-1)(1, 1) = (-1, -1) \notin W$$

(W is not closed under scalar multiplication)

$\therefore W$ is not a subspace of R^2

- Ex 6: Identify subspaces of R^2

Which of the following two subsets is a subspace of R^2?

(a) The set of points on the line given by $x + 2y = 0$

(b) The set of points on the line given by $x + 2y = 1$

Sol:

(a) $W = \{(x, y) \mid x + 2y = 0\} = \{(-2t, t) \mid t \in R\}$ (Note: the zero vector (0,0) is on this line)

Let $\mathbf{v}_1 = \left(-2t_1, t_1\right) \in W$ and $\mathbf{v}_2 = \left(-2t_2, t_2\right) \in W$

$\because \mathbf{v}_1 + \mathbf{v}_2 = \left(-2\left(t_1 + t_2\right), t_1 + t_2\right) \in W$ (closed under vector addition)

$c\mathbf{v}_1 = \left(-2\left(ct_1\right), ct_1\right) \in W$ (closed under scalar multiplication)

$\therefore W$ is a subspace of R^2

(b) $W=\{(x,y)\mid x+2y=1\}$ (Note: the zero vector (0, 0) is not on this line)

Consider $\mathbf{v}=(1,0)\in W$

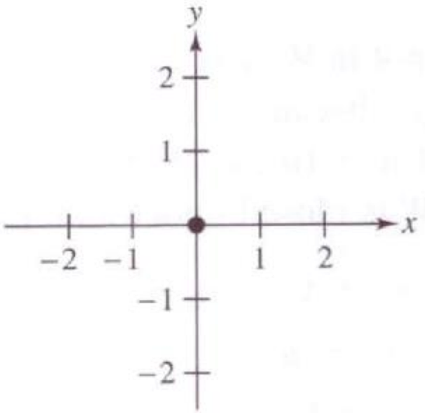
$\because(-1)\mathbf{v}=(-1,0)\notin W \quad \therefore W$ is not a subspace of R^2

- Notes: Subspaces of R^2

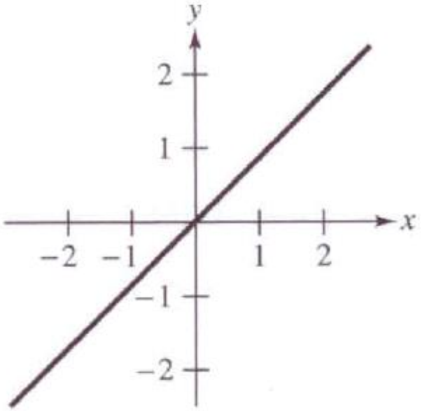
(1) W consists of the *single point* $\mathbf{0}=(0,0)$ (trivial subspace)

(2) W consists of all points on a *line* passing through the origin

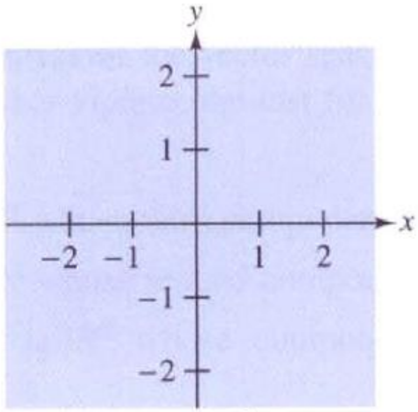
(3) R^2 (trivial subspace)



$W=\{(0,0)\}$



$W=$ all points on a line passing through the origin



$W=R^2$

- Ex 8: Identify subspaces of R^3

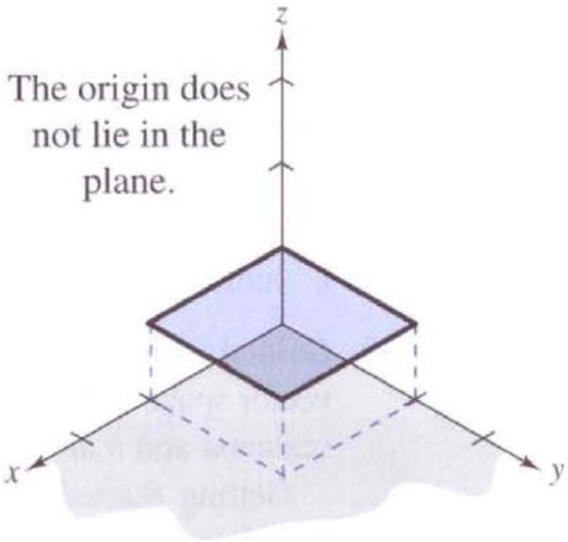
Which of the following subsets is a subspace of R^3?

(a) $W = \{(x_1, x_2, 1) \mid x_1, x_2 \in R\}$ (Note: the zero vector is not in W)

(b) $W = \{(x_1, x_1 + x_3, x_3) \mid x_1, x_3 \in R\}$ (Note: the zero vector is in W)

Sol:

(a)

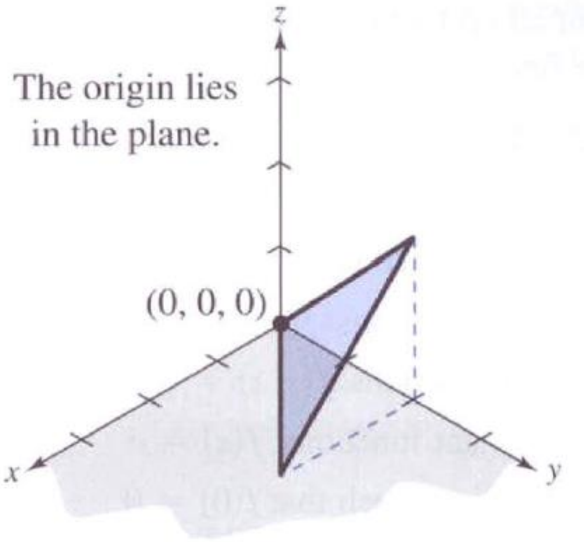


Consider $\mathbf{v} = (0, 0, 1) \in W$

$\because (-1)\mathbf{v} = (0, 0, -1) \notin W$

$\therefore W$ is not a subspace of R^3

(b)



Consider $\mathbf{v} = (v_1, v_1 + v_3, v_3) \in W$ and $\mathbf{u} = (u_1, u_1 + u_3, u_3) \in W$

$$\because \mathbf{v} + \mathbf{u} = \left(v_1 + u_1, \left(v_1 + u_1\right) + \left(v_3 + u_3\right), v_3 + u_3\right) \in W$$

$$c\mathbf{v} = \left(cv_1, \left(cv_1\right) + \left(cv_3\right), cv_3\right) \in W$$

$\therefore W$ is closed under vector addition and scalar multiplication,

so W is a subspace of R^3

- Notes: Subspaces of R^3

(1) W consists of the *single point* $\mathbf{0}=\left(0,0,0\right)$ (trivial subspace)

(2) W consists of all points on a *line* passing through the origin

(3) W consists of all points on a *plane* passing through the origin

(The W in problem (b) is a plane passing through the origin)

(4) R^3 (trivial subspace)

※ According to Ex. 6 and Ex. 8, we can infer that if W is a subspace of a vector space V, then both W and V must contain the same zero vector $\mathbf{0}$

- Note: The intersection of two subspaces is a subspace

If V and W are both subspaces of a vector space U,

then the intersection of V and W (denoted by $V \cap W$)

is also a subspace of U

However, the union of two subspaces is not a subspace. Prove that

Chapter 7

Eigenvalues and Eigenvectors

7.1 Eigenvalues and Eigenvectors

- Eigenvalue problem (one of the most important problems in the linear algebra):

  If A is an $n \times n$ matrix, do there exist nonzero vectors $\mathbf{x}$ in R^n such that $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$?

  (The term eigenvalue is from the German word *Eigenwert*, meaning "proper value")

- Eigenvalue and Eigenvector:

  A: an $n \times n$ matrix

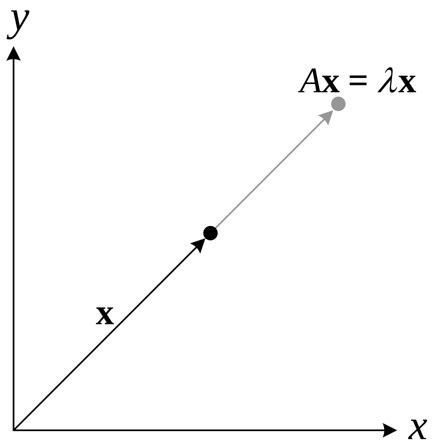
  λ: a scalar (could be **zero**)

  $\mathbf{x}$: a **nonzero** vector in R^n

$$A\underset{\uparrow}{\mathbf{x}} = \overset{\substack{\text{Eigenvalue}\\ \downarrow}}{\lambda}\underset{\uparrow}{\mathbf{x}}$$

Eigenvector

※ Geometric Interpretation



▪ Ex 1: Verifying eigenvalues and eigenvectors

$$A=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix}\quad \mathbf{x}_1=\begin{bmatrix}1\\ 0\end{bmatrix}\quad \mathbf{x}_2=\begin{bmatrix}0\\ 1\end{bmatrix}$$

$$A\mathbf{x}_1=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}1\\ 0\end{bmatrix}=\begin{bmatrix}2\\ 0\end{bmatrix}=\overset{\text{Eigenvalue}}{\overset{\downarrow}{2}}\underset{\underset{\text{Eigenvector}}{\uparrow}}{\begin{bmatrix}1\\ 0\end{bmatrix}}=2\mathbf{x}_1$$

$$A\mathbf{x}_2=\begin{bmatrix}2 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}0\\ 1\end{bmatrix}=\begin{bmatrix}0\\ -1\end{bmatrix}=\overset{\text{Eigenvalue}}{\overset{\downarrow}{-1}}\underset{\underset{\text{Eigenvector}}{\uparrow}}{\begin{bmatrix}0\\ 1\end{bmatrix}}=(-1)\mathbf{x}_2$$

※ In fact, for each eigenvalue, it has infinitely many eigenvectors. For $\lambda = 2$, $[3\ 0]^T$ or $[5\ 0]^T$ are both corresponding eigenvectors. Moreover, $([3\ 0] + [5\ 0])^T$ is still an eigenvector. The proof is in Thm. 7.1.

- Thm. 7.1: The eigenspace corresponding to λ of matrix A

  If A is an $n \times n$ matrix with an eigenvalue λ, then the set of all eigenvectors of λ **together with the zero vector** is a subspace of R^n. This subspace is called the eigenspace of λ

- Ex 3: Examples of eigenspaces on the *xy*-plane

  For the matrix A as follows, the corresponding eigenvalues are $\lambda_1 = -1$ and $\lambda_2 = 1$:

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Sol:

For the eigenvalue $\lambda_1 = -1$, corresponding vectors are any vectors on the *x*-axis

$$A\begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} -x \\ 0 \end{bmatrix} = \enclose{circle}{-1}\begin{bmatrix} x \\ 0 \end{bmatrix}$$

※ Thus, the eigenspace corresponding to $\lambda = -1$ is the x-axis, which is a subspace of R^2

For the eigenvalue $\lambda_2 = 1$, corresponding vectors are any vectors on the *y*-axis

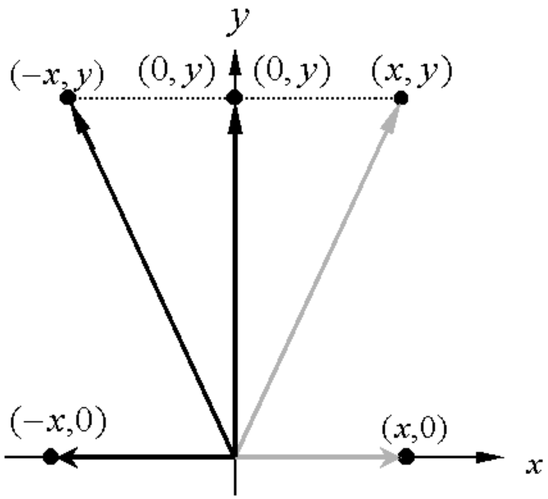
$$A\begin{bmatrix} 0 \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} = \enclose{circle}{1}\begin{bmatrix} 0 \\ y \end{bmatrix}$$

※ Thus, the eigenspace corresponding to $\lambda = 1$ is the y-axis, which is a subspace of R^2

※ Geometrically speaking, multiplying a vector (x, y) in R^2 by the matrix A corresponds to a reflection to the y-axis, i.e., left multiplying A to **v** can transform **v** to another vector in the same vector space

$$A\mathbf{v} = A\begin{bmatrix} x \\ y \end{bmatrix} = A\left(\begin{bmatrix} x \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ y \end{bmatrix}\right) = A\begin{bmatrix} x \\ 0 \end{bmatrix} + A\begin{bmatrix} 0 \\ y \end{bmatrix}$$

$$= -1\begin{bmatrix} x \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix}$$



- Thm. 7.2: Finding eigenvalues and eigenvectors of a matrix $A \in M_{n\times n}$

  Let A be an $n\times n$ matrix.

  (1) An eigenvalue of A is a scalar λ such that $\det(\lambda I - A) = 0$

  (2) The eigenvectors of A corresponding to λ are the nonzero solutions of $(\lambda I - A)\mathbf{x} = \mathbf{0}$

- Note: follwing the definition of the eigenvalue problem

  $A\mathbf{x} = \lambda\mathbf{x} \;\Rightarrow\; A\mathbf{x} = \lambda I\mathbf{x} \;\Rightarrow\; (\lambda I - A)\mathbf{x} = \mathbf{0}$ (homogeneous system)

  $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has nonzero solutions for $\mathbf{x}$ iff $\det(\lambda I - A) = 0$

  (The above iff results comes from the equivalent conditions on Slide 4.101)

- Characteristic equation of A:

  $$\det(\lambda I - A) = 0$$

- Characteristic polynomial of $A \in M_{n\times n}$:

  $$\det(\lambda I - A) = \left|(\lambda I - A)\right| = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

- Ex 4: Finding eigenvalues and eigenvectors

$$A = \begin{bmatrix} 2 & -12 \\ 1 & -5 \end{bmatrix}$$

Sol: Characteristic equation:

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 2 & 12 \\ -1 & \lambda + 5 \end{vmatrix}$$

$$= \lambda^2 + 3\lambda + 2 = (\lambda + 1)(\lambda + 2) = 0$$

$$\Rightarrow \lambda = -1, -2$$

Eigenvalue: $\lambda_1 = -1, \lambda_2 = -2$

$$(1)\ \lambda_1 = -1 \quad \Rightarrow (\lambda_1 I - A)\mathbf{x} = \begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -3 & 12 \\ -1 & 4 \end{bmatrix} \xrightarrow{\text{G.-J. E.}} \begin{bmatrix} 1 & -4 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4t \\ t \end{bmatrix} = t \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad t \neq 0$$

$$(2)\ \lambda_2 = -2 \quad \Rightarrow (\lambda_2 I - A)\mathbf{x} = \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -4 & 12 \\ -1 & 3 \end{bmatrix} \xrightarrow{\text{G.-J. E.}} \begin{bmatrix} 1 & -3 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3s \\ s \end{bmatrix} = s \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \quad s \neq 0$$

- Ex 5: Finding eigenvalues and eigenvectors

Find the eigenvalues and corresponding eigenvectors for the matrix A. What is the dimension of the eigenspace of each eigenvalue?

$$A=\begin{bmatrix}2 & 1 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{bmatrix}$$

Sol: Characteristic equation:

$$|\lambda I-A|=\begin{vmatrix}\lambda-2 & -1 & 0\\ 0 & \lambda-2 & 0\\ 0 & 0 & \lambda-2\end{vmatrix}=(\lambda-2)^3=0$$

Eigenvalue: $\lambda=2$

The eigenspace of $\lambda = 2$:

$$(\lambda I - A)\mathbf{x} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} s \\ 0 \\ t \end{bmatrix} = s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \ \ s, t \neq 0$$

$$\left\{ s\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \middle| s, t \in R \right\} : \text{the eigenspace of } A \text{ corresponding to } \lambda = 2$$

Thus, the dimension of its eigenspace is 2