Discrete Mathematics and Its Applications

Lecture 2: Basic Structures: Sequence, Cardinality, and Matrix

MING GAO

DaSE@ ECNU
(for course related communications)
mgao@dase.ecnu.edu.cn

Oct. 9, 2018

Outline

Sequence

Definition

A sequence is a function from a subset of the set of integers (usually either the set $\{0, 1, 2, \cdots\}$ or the set $\{1, 2, 3, \cdots\}$) to a set S. We use the notation a_n to denote the image of integer n. We call a_n a term of the sequence.

Sequence

Definition

A sequence is a function from a subset of the set of integers (usually either the set $\{0, 1, 2, \cdots\}$ or the set $\{1, 2, 3, \cdots\}$) to a set S. We use the notation a_n to denote the image of integer n. We call a_n a term of the sequence.

Some useful sequences

nth	Term First 10 Terms
n^2	$1, 4, 9, 16, 25, 36, 49, 64, 81, 100, \cdots$
n^3	$1, 8, 27, 64, 125, 216, 343, 512, 729, 1000, \cdots$
n^4	$1, 16, 81, 256, 625, 1296, 2401, 4096, 6561, 10000, \cdots$
$2n$	$2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, \cdots$
$3n$	$3, 9, 27, 81, 243, 729, 2187, 6561, 19683, 59049, \cdots$
$n!$	$1, 2, 6, 24, 120, 720, 5040, 40320, 362880, 3628800, \cdots$

Progression

Geometric progression

A geometric progression is a sequence of form $a, ar, ar^2, \cdots, ar^n, \cdots$ where initial term a and common ratio r are real numbers.

Progression

Geometric progression

A geometric progression is a sequence of form $a, ar, ar^2, \cdots, ar^n, \cdots$ where initial term a and common ratio r are real numbers.

Arithmetic progression

An arithmetic progression is a sequence of form $a, a+d, a+2d, \cdots, a+nd, \cdots$ where initial term a and common difference d are real numbers.

Progression

Geometric progression

A geometric progression is a sequence of form $a, ar, ar^2, \cdots, ar^n, \cdots$ where initial term a and common ratio r are real numbers.

Arithmetic progression

An arithmetic progression is a sequence of form $a, a+d, a+2d, \cdots, a+nd, \cdots$ where initial term a and common difference d are real numbers.

Examples

- The sequences $\{b_n\}$ is a form of $b_n = (-1)^n$. The list of terms $b_0, b_1, b_2, b_3, \cdots$ begins with $1, -1, 1, -1, \cdots$.
- The sequences $\{s_n\}$ is a form of $s_n = -1 + 4n$. The list of terms $s_0, s_1, s_2, s_3, \cdots$ begins with $-1, 3, 7, 11, \cdots$.

Recurrence relation

Definition

A recurrence relation for the sequence $\{a_n\}$ is an equation that expresses an in terms of one or more of the previous terms of the sequence, namely, $a_0, a_1, \cdots, a_{n-1}$, for all integers n with $n \geq n_0$, where n_0 is a nonnegative integer. A sequence is called a solution of a recurrence relation if its terms satisfy the recurrence relation.

Recurrence relation

Definition

A recurrence relation for the sequence $\{a_n\}$ is an equation that expresses an in terms of one or more of the previous terms of the sequence, namely, $a_0, a_1, \cdots, a_{n-1}$, for all integers n with $n \geq n_0$, where n_0 is a nonnegative integer. A sequence is called a solution of a recurrence relation if its terms satisfy the recurrence relation.

Example

- Let $\{a_n\}$ be a sequence that satisfies the recurrence relation $a_n = a_{n-1} + 3$ for $n = 1, 2, 3, \cdots$, and suppose that $a_0 = 2$. What are a_1, a_2, and a_3?
- Let $\{a_n\}$ be a sequence that satisfies the recurrence relation $a_n = a_{n-1} - a_{n-2}$ for $n = 2, 3, 4, \cdots$, and suppose that $a_0 = 3$ and $a_1 = 5$. What are a_2 and a_3?

Recurrence relation Cont'd

Fibonacci sequence

Fibonacci sequence, $f_0, f_1, f_2, \cdots$, is defined by initial conditions $f_0 = 0, f_1 = 1$, and recurrence relation $f_n = f_{n-1} + f_{n-2}$ for $n = 2, 3, 4, \cdots$

Recurrence relation Cont'd

Fibonacci sequence

Fibonacci sequence, $f_0, f_1, f_2, \cdots$, is defined by initial conditions $f_0 = 0, f_1 = 1$, and recurrence relation $f_n = f_{n-1} + f_{n-2}$ for $n = 2, 3, 4, \cdots$

Example

Determine whether sequence $\{a_n\}$, where $a_n = 3n$ for every nonnegative integer n, is a solution of recurrence relation $a_n = 2a_{n-1} - a_{n-2}$ for $n = 2, 3, 4, \cdots$ Answer the same question where $a_n = 2^n$ and where $a_n = 5$.

The Fibonacci sequence

Source:

https://en.wikipedia.org/wiki/

File:Fibonacci.jpg

In 1202, Leonardo Bonacci (known as Fibonacci) asked the following question.

> "[A]ssuming that: a newly born pair of rabbits, one male, one female, are put in a field; rabbits are able to mate at the age of one month so that at the end of its second month a female can produce another pair of rabbits; rabbits never die and a mating pair always produces one new pair (one male, one female) every month from the second month on."
>
> "The puzzle that Fibonacci posed was: how many pairs will there be in one year?"

From https://en.wikipedia.org/wiki/Fibonacci_number

Let's try to solve Fibonacci's question.

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡	

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡	

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3
5	♡ ♡ ♡	

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3
5	♡ ♡ ♡ ♠ ♠	5

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3
5	♡ ♡ ♡ ♠ ♠	5
6	♡ ♡ ♡ ♡ ♡	

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3
5	♡ ♡ ♡ ♠ ♠	5
6	♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠	8

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3
5	♡ ♡ ♡ ♠ ♠	5
6	♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠	8
7	♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡	

Let's try to solve Fibonacci's question.

Let $\spadesuit$ denote a newly born rabit pair, and $\heartsuit$ denote a mature rabit pair.

Month	Rabits	
1	$\spadesuit$	1
2	$\heartsuit$	1
3	$\heartsuit$ $\spadesuit$	2
4	$\heartsuit$ $\heartsuit$ $\spadesuit$	3
5	$\heartsuit$ $\heartsuit$ $\heartsuit$ $\spadesuit$ $\spadesuit$	5
6	$\heartsuit$ $\heartsuit$ $\heartsuit$ $\heartsuit$ $\heartsuit$ $\spadesuit$ $\spadesuit$ $\spadesuit$	8
7	$\heartsuit$ $\heartsuit$ $\heartsuit$ $\heartsuit$ $\heartsuit$ $\heartsuit$ $\heartsuit$ $\heartsuit$ $\spadesuit$ $\spadesuit$ $\spadesuit$ $\spadesuit$ $\spadesuit$	13

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3
5	♡ ♡ ♡ ♠ ♠	5
6	♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠	8
7	♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠ ♠ ♠	13

How many rabit pairs do we have at the beginning of the 8th month?

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3
5	♡ ♡ ♡ ♠ ♠	5
6	♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠	8
7	♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠ ♠ ♠	13

How many rabit pairs do we have at the beginning of the 8th month?

- Surely all 13 rabit pairs we have in the 7th month remain there and are all mature. So, the question is how many newly born rabbit pairs that we have.

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3
5	♡ ♡ ♡ ♠ ♠	5
6	♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠	8
7	♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠ ♠ ♠	13

How many rabit pairs do we have at the beginning of the 8th month?

- Surely all 13 rabit pairs we have in the 7th month remain there and are all mature. So, the question is how many newly born rabbit pairs that we have.
- The number of newly born rabbit pairs equals the number of mature rabbit pairs we have.

Let's try to solve Fibonacci's question.

Let ♠ denote a newly born rabit pair, and ♡ denote a mature rabit pair.

Month	Rabits	
1	♠	1
2	♡	1
3	♡ ♠	2
4	♡ ♡ ♠	3
5	♡ ♡ ♡ ♠ ♠	5
6	♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠	8
7	♡ ♡ ♡ ♡ ♡ ♡ ♡ ♡ ♠ ♠ ♠ ♠ ♠	13

How many rabit pairs do we have at the beginning of the 8th month?

- Surely all 13 rabit pairs we have in the 7th month remain there and are all mature. So, the question is how many newly born rabbit pairs that we have.
- The number of newly born rabbit pairs equals the number of mature rabbit pairs we have. This is also equal to the number of rabit pairs that we have in the 6th month: 8.

Thus, we will have 13+8 rabit pairs at the beginning of the 8th month.

Thus, we will have 13+8 rabit pairs at the beginning of the 8th month. If we write down the sequence, we get the Fibonacci sequence:

$$1, 1, 2, 3, 5, 8, 13, 21, \ldots$$

Thus, we will have 13+8 rabit pairs at the beginning of the 8th month.
If we write down the sequence, we get the Fibonacci sequence:

$$1, 1, 2, 3, 5, 8, 13, 21, \ldots$$

Again, what's the next number in this sequence? How can you compute it?

Thus, we will have 13+8 rabit pairs at the beginning of the 8th month.
If we write down the sequence, we get the Fibonacci sequence:

$$1, 1, 2, 3, 5, 8, 13, 21, \ldots$$

Again, what's the next number in this sequence? How can you compute it?
$21+13 = 34$ is the answer.

Thus, we will have 13+8 rabit pairs at the beginning of the 8th month. If we write down the sequence, we get the Fibonacci sequence:

$$1, 1, 2, 3, 5, 8, 13, 21, \ldots$$

Again, what's the next number in this sequence? How can you compute it? $21+13 = 34$ is the answer. You take the last two numbers and add them up to get the next number. Why?

To be precise, let F_n be the n-th number in the Fibonacci sequence. (That is, $F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3$ and so on.) We can define the $(n+1)$-th number as

$$F_{n+1} = F_n + F_{n-1},$$

for $n = 2, 3, \ldots$.

To be precise, let F_n be the n-th number in the Fibonacci sequence. (That is, $F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3$ and so on.) We can define the $(n+1)$-th number as

$$F_{n+1} = F_n + F_{n-1},$$

for $n = 2, 3, \ldots$. Is this enough to completely specify the sequence?

To be precise, let F_n be the n-th number in the Fibonacci sequence. (That is, $F_1 = 1, F_2 = 1, F_3 = 2, F_4 = 3$ and so on.) We can define the $(n+1)$-th number as

$$F_{n+1} = F_n + F_{n-1},$$

for $n = 2, 3, \ldots$. Is this enough to completely specify the sequence? No, because we do not know how to start. To get the Fibonacci sequence, we need to specify two starting values: $F_1 = 1$ and $F_2 = 1$ as well. Now, you can see that the equation and these special values uniquely determine the sequence. It is also convenient to define $F_0 = 0$ so that the equation works for $n = 1$.

A recurrence

The equation

$$F_{n+1} = F_n + F_{n-1}$$

and the initial values $F_0 = 0$ and $F_1 = 1$ specify all values of the Fibonacci sequence. With these two initial values, you can use the equation to find the value of any number in the sequence.

This definition is called a **recurrence**. Instead of defining the value of each number in the sequence explicitly, we do so by using the values of other numbers in the sequence.

Tilings with 1x1 and 2x1 tiles

You have a walk way of length n units. The width of the walk way is 1 unit. You have unlimited supplies of 1x1 tiles and 2x1 tiles. Every tile of the same size is indistinguishable. In how many ways can you tile the walk way?

Let's consider small cases.

- When $n = 1$, there are 1 way.
- When $n = 2$, there are 2 ways.
- When $n = 3$, there are 3 ways.
- When $n = 4$, there are 5 ways.

Let's define J_n to be the number of ways you can tile a walk way of length n. From the example above, we know that $J_1 = 1$ and $J_2 = 2$.

Can you find a formula for general J_n?

Figuring out the recurrence for J_n

To figure out the general formula for J_n, we can think about the first choice we can make when tiling a walk way of length n. There are two choices:

- (1) We can start placing a 1x1 tile at the beginning, or
- (2) We can start placing a 2x1 tile at the beginning.

In each of the cases, let's think about how many ways we can tile the rest of the walk way, provided that the first step is made.

Note that if we start by placing a 1x1 tile, we are left with a walk way of length $n-1$. From the definition of J_n, we know that there are J_{n-1} ways to tile the rest of the walk way of length $n-1$. Using similar reasoning, we know that if we start with a 2x1 tile, there are J_{n-2} ways to tile the rest of the walk way.

The recurrence for J_n

From the discussion, we have that

$$J_n = J_{n-1} + J_{n-2},$$

where $J_1 = 1$ and $J_2 = 2$.

Note that this is exactly the same recurrence as the Fibonacci sequence, but with different initial values. In fact, we have that

$$J_n = F_{n+1}.$$

Compound interest

Problem

Suppose that a person deposits $10,000$ in a savings account at a bank yielding 11% per year with interest compounded annually. How much will be in the account after 30 years?

Compound interest

Problem

Suppose that a person deposits $10,000$ in a savings account at a bank yielding 11% per year with interest compounded annually. How much will be in the account after 30 years?

Solution

Let P_n denote the amount in the account after n years. We see that sequence $\{P_n\}$ satisfies the recurrence relation

$$P_n = P_{n-1} + 0.11P_{n-1} = (1.11)P_{n-1}.$$

We can find a formula for P_n:

$$P_1 = (1.11)P_0 \quad P_2 = (1.11)P_1 = (1.11)^2 P_0$$
$$\cdots \quad P_n = (1.11)P_{n-1} = \cdots = (1.11)^n P_0.$$

Therefore, $P_{30} = (1.11)^{30} 10,000 = 228,922.97.$

Summations

Summation notations

$$\sum_{j=m}^{n} a_j, \quad \sum_{i=m}^{n} a_i, \quad \sum_{m\le k\le n} a_k, \quad \sum_{j\in I} a_j,$$

where $I = \{k \in Z : m \le k \le n\}$.

Summations

Summation notations

$$\sum_{j=m}^{n} a_j, \quad \sum_{i=m}^{n} a_i, \quad \sum_{m\leq k\leq n} a_k, \quad \sum_{j\in I} a_j,$$

where $I = \{k \in Z : m \leq k \leq n\}$.

Some useful summation formulae

sum	form	sum	form
$\sum_{k=0}^{n} ar^k$	$\frac{ar^{n+1}-a}{r-1}(r \neq 1)$	$\sum_{k=0}^{n} k$	$\frac{n(n+1)}{2}$
$\sum_{k=0}^{n} k^2$	$\frac{n(n+1)(2n+1)}{6}$	$\sum_{k=0}^{n} k^3$	$\frac{n^2(n+1)^2}{4}$
$\sum_{k=0}^{\infty} x^k, \lvert x\rvert < 1$	$\frac{1}{1-x}$	$\sum_{k=0}^{\infty} kx^{k-1}, \lvert x\rvert < 1$	$\frac{1}{(1-x)^2}$

Remark

Requires calculus

Let x be a real number with $|x| < 1$. Find

$$\sum_{n=0}^{\infty} x^n.$$

Solution

Let $a = 1$ and $r = x$ we see that $\sum_{k=0}^{n} x^k = \frac{x^{k+1}-1}{x-1}$. Because $|x| < 1$, it follows that

$$\sum_{k=0}^{\infty} x^n = \lim_{k\to\infty} \frac{x^{k+1}-1}{x-1} = \frac{0-1}{x-1} = \frac{1}{1-x}.$$

Cardinality

Definition

Sets A and B have the same cardinality if and only if there is a one-to-one correspondence from A to B. When A and B have the same cardinality, we write $|A| = |B|$.

Cardinality

Definition

Sets A and B have the same cardinality if and only if there is a one-to-one correspondence from A to B. When A and B have the same cardinality, we write $|A| = |B|$.

Definition

If there is a one-to-one function from A to B, the cardinality of A is less than or the same as the cardinality of B and we write $|A| \leq |B|$. Moreover, when $|A| \leq |B|$ and A and B have different cardinality, we say that the cardinality of A is less than the cardinality of B and we write $|A| < |B|$.

Countable sets

Definition

A set that is either finite or has the same cardinality as Z^+ is called countable. If an infinite set S is countable, then $|S| = \aleph_0$.

Countable sets

Definition

A set that is either finite or has the same cardinality as Z^+ is called countable. If an infinite set S is countable, then $|S| = \aleph_0$.

Example

Show that the set of odd positive integers is a countable set.

Solution

In terms of the same cardinality, we will exhibit a one-to-one correspondence between this set and Z^+. Consider function

$$f(n) = 2n - 1 : Z^+ \to \{2k + 1 : k \text{ is an integer}\}$$

Countable sets

Definition

A set that is either finite or has the same cardinality as Z^+ is called countable. If an infinite set S is countable, then $|S| = \aleph_0$.

Example

Show that the set of odd positive integers is a countable set.

Solution

In terms of the same cardinality, we will exhibit a one-to-one correspondence between this set and Z^+. Consider function

$$f(n) = 2n - 1 : Z^+ \to \{2k + 1 : k \text{ is an integer}\}$$

One-to-one: $f(n) = f(m) \Rightarrow 2n - 1 = 2m - 1 \Rightarrow n = m.$

Countable sets

Definition

A set that is either finite or has the same cardinality as Z^+ is called countable. If an infinite set S is countable, then $|S| = \aleph_0$.

Example

Show that the set of odd positive integers is a countable set.

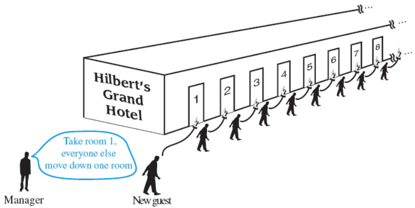
Solution

In terms of the same cardinality, we will exhibit a one-to-one correspondence between this set and Z^+. Consider function

$$f(n) = 2n - 1 : Z^+ \rightarrow \{2k + 1 : k \textit{ is an integer}\}$$

One-to-one: $f(n) = f(m) \Rightarrow 2n - 1 = 2m - 1 \Rightarrow n = m$.
Onto: suppose that t is an odd positive integer. Then t is 1 less than $2k$, where k is a natural number. Hence $t = 2k - 1 = f(k)$.

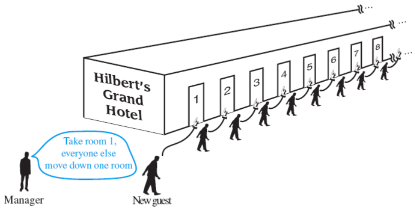
Hilberts Grand Hotel



Grand Hotel with $\aleph_0$ rooms

How can we accommodate a new guest arriving at the fully occupied Grand Hotel without removing any of the current guests?

Hilberts Grand Hotel



Grand Hotel with $\aleph_0$ rooms

How can we accommodate a new guest arriving at the fully occupied Grand Hotel without removing any of the current guests?

Solution

Because the rooms of the Grand Hotel are countable, we can list them as Room 1, Room 2, Room 3, and so on. When a new guest arrives, we move the guest in Room 1 to Room 2, the guest in Room 2 to Room 3, and in general, the guest in Room n to Room $n+1$, for all positive integers n. This frees up Room 1, which we assign to the new guest, and all the current guests still have rooms.

Countable sets

Integer set

Show that the set of all integers is countable.

Solution

$$f(n) = \begin{cases} n/2, & n \text{ is an even number;} \\ -(n-1)/2, & \text{otherwise.} \end{cases}$$

Countable sets

Integer set

Show that the set of all integers is countable.

Solution

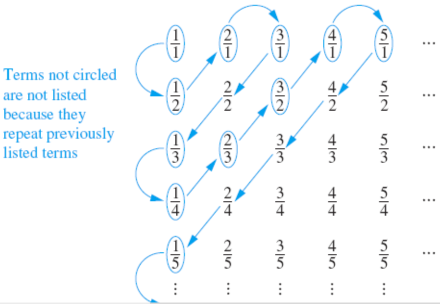
$$f(n) = \begin{cases} n/2, & n \textit{ is an even number;} \\ -(n-1)/2, & \textit{otherwise.} \end{cases}$$

Positive rational numbers

Show that all positive rational numbers is countable.

Solution

A rational numbers can be listed in p/q first with $p+q=2$, then $p+q=3, \cdots$



An uncountable set

Real numbers

Show that the set of real numbers is an uncountable set.

An uncountable set

Real numbers

Show that the set of real numbers is an uncountable set.

Solution

We suppose that the set of real numbers is countable and arrive at a contradiction.

An uncountable set

Real numbers

Show that the set of real numbers is an uncountable set.

Solution

We suppose that the set of real numbers is countable and arrive at a contradiction.

Then, $\forall x \in (0,1)$ would also be countable (because any subset of a countable set is also countable).

Under this assumption, $\forall x \in (0,1)$ can be listed in some order, say, $r_1, r_2, r_3, \cdots$ Let the decimal representation of these real numbers be

$$r_1 = 0.d_{11}d_{12}d_{13}d_{14}\cdots$$
$$r_2 = 0.d_{21}d_{22}d_{23}d_{24}\cdots$$
$$r_3 = 0.d_{31}d_{32}d_{33}d_{34}\cdots$$
$$\cdots$$

An uncountable set Cont'd

Solution

For example, if $r_1 = 0.23794\cdots$, we have $d_{11} = 2$, $d_{12} = 3$, etc.)

An uncountable set Cont'd

Solution

For example, if $r_1 = 0.23794\cdots$, we have $d_{11} = 2$, $d_{12} = 3$, etc.)
Then, form a new real number $r = 0.d_1d_2d_3d_4\cdots$, where the decimal digits are determined by the following rule:

$$d_i = \begin{cases} 4, & \text{if } d_{ii} \neq 4; \\ 5, & \text{otherwise.} \end{cases}$$

For example, let $r_1 = 0.23794\cdots, r_2 = 0.44590\cdots, r_3 = 0.09118\cdots, r_4 = 0.80553\cdots$, etc. Then we have $r = 0.d_1d_2d_3d_4\cdots = 0.4544\cdots$, where $d_1 = 4$ since $d_{11} \neq 4, d_2 = 5$ since $d_{22} = 4, d_3 = 4$ since $d_{33} \neq 4, d_4 = 4$ since $d_{44} \neq 4$, etc.

An uncountable set Cont'd

Solution

For example, if $r_1 = 0.23794\cdots$, we have $d_{11} = 2$, $d_{12} = 3$, etc.)
Then, form a new real number $r = 0.d_1d_2d_3d_4\cdots$, where the decimal digits are determined by the following rule:

$$d_i = \begin{cases} 4, & \text{if } d_{ii} \neq 4; \\ 5, & \text{otherwise.} \end{cases}$$

For example, let $r_1 = 0.23794\cdots, r_2 = 0.44590\cdots, r_3 = 0.09118\cdots, r_4 = 0.80553\cdots$, etc. Then we have $r = 0.d_1d_2d_3d_4\cdots = 0.4544\cdots$, where $d_1 = 4$ since $d_{11} \neq 4, d_2 = 5$ since $d_{22} = 4, d_3 = 4$ since $d_{33} \neq 4, d_4 = 4$ since $d_{44} \neq 4$, etc.
Every real number has a unique decimal expansion. Therefore, r is not equal to any of $r_1, r_2, \cdots$ because the decimal expansion of r differs from the decimal expansion of r_i in the i-th place to the right of the decimal point, for each i.

Results about cardinality

Theorem

If A and B are countable sets, then $A \cup B$ is also countable.

Results about cardinality

Theorem

If A and B are countable sets, then $A \cup B$ is also countable.

Proof.

Without loss of generality, we can assume that A and B are disjoint. (If they are not, we can replace B by $B - A$, because $A \cap (B - A) = \emptyset$ and $A \cup (B - A) = A \cup B$.)

Results about cardinality

Theorem

If A and B are countable sets, then $A \cup B$ is also countable.

Proof.

Without loss of generality, we can assume that A and B are disjoint. (If they are not, we can replace B by $B - A$, because $A \cap (B - A) = \emptyset$ and $A \cup (B - A) = A \cup B$.)

C1: A and B are both finite;

C2: A is infinite and B is finite;

C3: A and B are both countably infinite.

□

SCHRÖDER-BERNSTEIN theorem and the continuum hypothesis

Theorem

If A and B are sets with $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$, i.e., if there are one-to-one functions f from A to B and g from B to A, then there is a one-to-one correspondence between A and B.

SCHRÖDER-BERNSTEIN theorem and the continuum hypothesis

Theorem

If A and B are sets with $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$, i.e., if there are one-to-one functions f from A to B and g from B to A, then there is a one-to-one correspondence between A and B.

Hypothesis

Show that the power set of Z^+ and R have the same cardinality, i.e., $|\mathcal{P}(Z^+)| = |R| = c$.
Furthermore, $|P(Z^+)| = |R|$ can be expressed as $2^{\aleph_0} = c$.

SCHRÖDER-BERNSTEIN theorem and the continuum hypothesis

Theorem

If A and B are sets with $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$, i.e., if there are one-to-one functions f from A to B and g from B to A, then there is a one-to-one correspondence between A and B.

Hypothesis

Show that the power set of Z^+ and R have the same cardinality, i.e., $|\mathcal{P}(Z^+)| = |R| = c$.
Furthermore, $|P(Z^+)| = |R|$ can be expressed as $2^{\aleph_0} = c$.

Cantor theorem

The cardinality of a set is always less than the cardinality of its power set, i.e., $|Z^+| < |\mathcal{P}(Z^+)|$.

Linear equations and matrix

Linear equations

- The subject of algebra arose from studying equations.
- If $x_1, x_2, \cdots, x_n$ are variables and $a_1, a_2, \cdots, a_n$ and c are constant, then the equation $a_1x_1 + a_2x_2 + \cdots + a_nx_n = c$ is said to be a linear equation, where a_i are the coefficient.
- More generally, a linear system consisting of m equations in n unknowns will look like:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= c_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= c_2 \\ &\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= c_m \end{aligned}$$

- The main problem is to find the solution set of a linear system (Gaussian reduction). While that is not the focus of this course

Vector

An n-tuple (pair, triple, quadruple, ...) of scalars can be written as a horizontal row or vertical column. A column is called a vector. For example $x = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$ and $x^T = [x_1, x_2, \cdots, x_n]$

Vector

An n-tuple (pair, triple, quadruple, ...) of scalars can be written as a horizontal row or vertical column. A column is called a vector. For example $x = \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$ and $x^T = [x_1, x_2, \cdots, x_n]$

Operations

- Addition: $x + y = [x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n]$
  - $x + y = y + x$
  - $cx = [cx_1, cx_2, \cdots, cx_n]$ and $0x = \overrightarrow{0}$
  - $x + (y + z) = (x + y) + z$
- Manipulation: $x^T y = \sum_{i=1}^{n} x_i y_i$
  - $x^T y = y^T x$
  - $(x + y)^T z = x^T z + y^T z$

Examples of vector

Examples

- Entity: an entity can be modeled as a vector $x = [x_1, x_2, \cdots, x_n]$, where n denotes the number of features and x_i denotes the value of the $i-$th feature. For example, patients, email, student, and user, etc.

Examples of vector

Examples

- Entity: an entity can be modeled as a vector $x = [x_1, x_2, \cdots, x_n]$, where n denotes the number of features and x_i denotes the value of the $i-$th feature. For example, patients, email, student, and user, etc.
- Set: given a universal set, a subset of the universal set can be model as a binary vector $x = [0, 1, 1, \cdots, 0]^T$, where the dimension of X is the size of universal set and $x_i = 1$ means that the $i-$th item exists in the set. For example, document VS. words, vertex VS. neighbors, string VS. n-gram, user VS. products, etc.

Examples of vector

Examples

- Entity: an entity can be modeled as a vector $x = [x_1, x_2, \cdots, x_n]$, where n denotes the number of features and x_i denotes the value of the $i-$th feature. For example, patients, email, student, and user, etc.
- Set: given a universal set, a subset of the universal set can be model as a binary vector $x = [0, 1, 1, \cdots, 0]^T$, where the dimension of X is the size of universal set and $x_i = 1$ means that the $i-$th item exists in the set. For example, document VS. words, vertex VS. neighbors, string VS. n-gram, user VS. products, etc.
- Distribution: given a discrete sample space Ω, PMF can be modeled as a vector $p = [p_1, p_2, \cdots, p_n]$, where n is the cardinality of Ω, $0 \le p_i \le 1$ and $\sum_{i=1}^{n} p_i = 1$. For example, document VS. topics, Markov chain VS. states, tweet VS. polarity, entity VS. classes, etc.

Examples of vector

Examples

- Entity: an entity can be modeled as a vector $x = [x_1, x_2, \cdots, x_n]$, where n denotes the number of features and x_i denotes the value of the $i-$th feature. For example, patients, email, student, and user, etc.
- Set: given a universal set, a subset of the universal set can be model as a binary vector $x = [0, 1, 1, \cdots, 0]^T$, where the dimension of X is the size of universal set and $x_i = 1$ means that the $i-$th item exists in the set. For example, document VS. words, vertex VS. neighbors, string VS. n-gram, user VS. products, etc.
- Distribution: given a discrete sample space Ω, PMF can be modeled as a vector $p = [p_1, p_2, \cdots, p_n]$, where n is the cardinality of Ω, $0 \leq p_i \leq 1$ and $\sum_{i=1}^{n} p_i = 1$. For example, document VS. topics, Markov chain VS. states, tweet VS. polarity, entity VS. classes, etc.
- Latent vector: matrix factorization, topic modeling, and word embedding, etc.

How far from two vectors

Distance

Distance is a numerical description of how far apart vectors are. Given two vectors x and y, there are many ways to measure distance of two vectors.

- Distance in Euclidean space: the Minkowski distance of order p (p-norm distance) is defined as
    - 1-norm distance (Manhattan distance): $\sum_{i=1}^{n} |x_i - y_i|$
    - 2-norm distance (Euclidean distance): $\left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}$
    - p-norm distance: $\left(\sum_{i=1}^{n} (x_i - y_i)^p\right)^{1/p}$
    - Infinity norm distance: $\lim_{p\to\infty} \left(\sum_{i=1}^{n} (x_i - y_i)^p\right)^{1/p} = \max\{x_i - y_i | i = 1, 2, \cdots, n\}$
- Mahalanobis distance: it is defined as a dissimilarity between two random vectors X and Y of the same distribution with covariance matrix S: $D(X, Y) = \left((X - Y)^T S^{-1} (X - Y)\right)^{1/2}$. If $S = I$, the Mahalanobis distance reduces to the Euclidean distance

Matrix

Definition

An $m \times n$ matrix $A = (a_{ij})$ $(1 \le i \le m, 1 \le j \le n)$ is a rectangular array of mn scalars in m rows and n columns, such as

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

The i-th row of A is the $1 \times n$ matrix $[a_{i1}, a_{i2}, \cdots, a_{in}]$. The $j-$th column of A is the $m \times 1$ matrix $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \cdots \\ a_{mj} \end{pmatrix}$.

Matrix

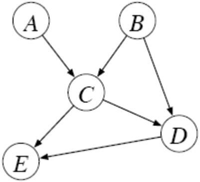
Concepts

- Boldface uppercase letters will be used to represent matrices;
- The set of all $m \times n$ matrices with real entries will be denoted by $\mathbb{R}^{m \times n}$;
- A is called the identity matrix if

$$a_{ij} = \begin{cases} 1, & \text{if } i = j; \\ 0, & \text{otherwise.} \end{cases}$$

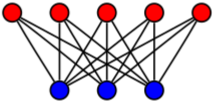
- If $m = n$, A is called a square matrix;
- $A^T = (a_{ji})(\in \mathbb{R}^{n \times m})$ is a $n \times m$ matrix;
- If A is a $n \times n$ square matrix, $Trace(A) = \sum_i^n a_{ii} = Trace(A^T)$;
- $AI_n = I_m A = A$;
- A square matrix A is called symmetric if $A = A^T$, i.e., $a_{ij} = a_{ji}$ for $\forall i$ and $\forall j$.

Examples of matrix



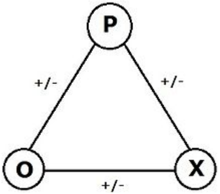
Homogeneous graph

Vertex can be Web page, user, protein, road, route, etc. Edge can be directed, undirected, weighted or labeled.



Heterogeneous graph

Vertex can be user-Web page, user-product, user-service, user-paper, etc.



Signed graph

user friend and enemy networks, like and dislike networks, etc.

Examples of matrix Cont.

Image processing

color, texture and shape, etc

Object	F1	F2	F3	F4
1	1.3	12.5	0.4	234.8
2	2.2	23.1	0.45	255.6
3	1.9	7.4	0.54	301.3
4	?	14.2	0.51	278.3

Feature representation

Many applications can be modeled in this manner, such as search engine, email classification, disease diagnosis, churn predication, anomaly detection, etc.

Operations of matrix

Operations

- Addition: Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $m \times n$ matrices, $A + B = (a_{ij} + b_{ij})$
  - $A + cB = (a_{ij} + cb_{ij})$, especially $A - B = (a_{ij} - b_{ij})$, where c is a scalar
  - $A + B = B + A$
  - $A + (B + C) = (A + B) + C$
- Manipulation: If A is an $n \times m$ matrix and B is an $m \times p$ matrix, then $AB = (\sum_{k=1}^{m} a_{ik} b_{kj})$
  - Not commutative: $AB \neq BA$
  - Distributive over matrix addition: $(A + B)C = AC + BC$
  - Scalar multiplication is compatible with matrix multiplication: $\lambda AB = (\lambda A)B = A(\lambda B)$
  - $(AB)^T = B^T A^T$
  - $Trace(AB) = Trace(BA)$ and $Trace(ABCD) = Trace(BCDA) = Trace(CDAB) = Trace(DABC)$. In general, $Trace(ABC) \neq Trace(ACB)$

Power of matrices

Definition

Let A be an $n \times n$ matrix, we have

$$A^0 = I_n, A^k = \underbrace{AA \cdots A}_{k \text{ times}}$$

Power of matrices

Definition

Let A be an $n \times n$ matrix, we have

$$A^0 = I_n, A^k = \underbrace{AA \cdots A}_{k \text{ times}}$$

Definition of diagonalizable

A square matrix A is said to be diagonalizable if it is similar to a diagonal matrix, i.e., there exists an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

- Why useful? If A is diagonalizable, then $A^k = PD^kP^{-1}$ for $k > 0$.
- How to find diagonal matrix? If $v_1, \cdots, v_n$ are linearly independent eigenvectors of A and λ_i are their corresponding eignevalues, then $A = PDP^{-1}$, where $P = [v_1 \ \cdots \ v_n]$ and $D = Diag(\lambda_1, \cdots, \lambda_n)$.

Matrix inverse

Definition

Suppose two $n \times n$ matrices A and B have the property that $AB = BA = I_n$. Then we say A is an inverse of B (and B is an inverse of A). Matrix A is invertible if A has inverse.

Properties

- Suppose $A \in \mathbb{R}^{n\times n}$ and A has an inverse B. Then B is unique.
- $A \in \mathbb{R}^{n\times n}$ is nonsingular (i.e., has rank n), if and only if there exists a matrix $B \in \mathbb{R}^{n\times n}$ such that $BA = I_n$.
- If A is nonsingular, then the system $Ax = b$ has the unique solution $x = A^{-1}b$.
- $(A|I_n) \to (I_n|B)$ (row reduce), then B is the inverse of A.
- $(A|\mathbf{b}) \to (I_n|\mathbf{c})$ (row reduce), where the components of $\mathbf{c}$ are certain linear combinations of the components of $\mathbf{b}$. The coefficients in these linear combinations give us the entries of A^{-1}.

Matrix derivatives

Type	scalar	vector	matrix
scalar	$\frac{\partial y}{\partial x}$	$\frac{\partial \mathbf{y}}{\partial x}$	$\frac{\partial Y}{\partial x}$
vector	$\frac{\partial y}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$	
matrix	$\frac{\partial y}{\partial X}$		

Matrix derivatives

Type	scalar	vector	matrix
scalar	$\frac{\partial y}{\partial x}$	$\frac{\partial \mathbf{y}}{\partial x}$	$\frac{\partial Y}{\partial x}$
vector	$\frac{\partial y}{\partial \mathbf{x}}$	$\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$	
matrix	$\frac{\partial y}{\partial X}$		

Derivatives by scalar

Assume that $x, y, a \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^{n\times 1}$, $X, Y, A \in \mathbb{R}^{n\times m}$, and $\mathbf{y}, \mathbf{a} \in \mathbb{R}^{m\times 1}$. Let a, $\mathbf{a}$ and A be constant scalar, vector and matrix.

- $\frac{\partial y}{\partial x}$ and $\frac{\partial a}{\partial x} = 0$
- $\frac{\partial \mathbf{y}}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \cdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}$ and $\frac{\partial \mathbf{a}}{\partial x} = \mathbf{0}$ (vector)
- $\frac{\partial Y}{\partial x} = \begin{bmatrix} \frac{\partial y_{11}}{\partial x} & \cdots & \frac{\partial y_{1m}}{\partial x} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_{n1}}{\partial x} & \cdots & \frac{\partial y_{nm}}{\partial x} \end{bmatrix}$ and $\frac{\partial A}{\partial x} = \mathbf{0}$ (matrix)

Matrix derivatives Cont.

Derivatives by vector

- $\frac{\partial y}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \cdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}^T$ and $\frac{\partial a}{\partial \mathbf{x}} = \mathbf{0}^T$ (vector)
- $\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$, $\frac{\partial \mathbf{a}}{\partial \mathbf{x}} = \mathbf{0}$ (matr.) and $\frac{\partial \mathbf{x}}{\partial \mathbf{x}} = \mathbf{I}$ (matr.)

Matrix derivatives Cont.

Derivatives by vector

- $\frac{\partial y}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \cdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix}^T$ and $\frac{\partial a}{\partial \mathbf{x}} = \mathbf{0}^T$ (vector)
- $\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$, $\frac{\partial \mathbf{a}}{\partial \mathbf{x}} = \mathbf{0}$ (matr.) and $\frac{\partial \mathbf{x}}{\partial \mathbf{x}} = \mathbf{I}$ (matr.)

Derivatives by matrix

$$\frac{\partial y}{\partial X} = \begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \cdots & \frac{\partial y}{\partial x_{n1}} \\ \cdots & \cdots & \cdots \\ \frac{\partial y}{\partial x_{1m}} & \cdots & \frac{\partial y}{\partial x_{nm}} \end{bmatrix} \text{ and } \frac{\partial a}{\partial X} = \mathbf{0}^T \text{ (matrix)}$$

Common properties of matrix derivatives

Properties

- c1 $\frac{\partial \mathbf{a}^T\mathbf{x}}{\partial \mathbf{x}} = \frac{\partial \mathbf{x}^T\mathbf{a}}{\partial \mathbf{x}} = \mathbf{a}^T$
- c2 $\frac{\partial \mathbf{x}^T\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{x}^T$
- c3 $\frac{\partial (\mathbf{x}^T\mathbf{a})^2}{\partial \mathbf{x}} = 2\mathbf{x}^T\mathbf{a}\mathbf{a}^T$
- c4 $\frac{\partial A\mathbf{x}}{\partial \mathbf{x}} = A$ and $\frac{\partial \mathbf{x}^T A}{\partial \mathbf{x}} = A^T$
- c5 $\frac{\partial \mathbf{x}^T A\mathbf{x}}{\partial \mathbf{x}} = \mathbf{x}^T(A + A^T)$

Proof c2. Let $s = \mathbf{x}^T\mathbf{x} = \sum_{i=1}^n x_i^2$. Then, $\frac{\partial s}{\partial x_i} = 2x_i$. So, $\frac{\partial \mathbf{x}^T\mathbf{x}}{\partial \mathbf{x}} = 2\mathbf{x}^T$.

c3. Let $s = \mathbf{x}^T\mathbf{a}$. Then $\frac{\partial s^2}{\partial x_i} = 2s\frac{\partial s}{\partial x_i} = 2sa_i$. Thus, $\frac{\partial (\mathbf{x}^T\mathbf{a})^2}{\partial \mathbf{x}} = 2\mathbf{x}^T\mathbf{a}\mathbf{a}^T$.

Properties of matrix derivatives cont.

Properties for scalar by scaler

ss1 $\frac{\partial(u+v)}{\partial x} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial x}$

ss2 $\frac{\partial uv}{\partial x} = u\frac{\partial v}{\partial x} + v\frac{\partial u}{\partial x}$ (product rule)

ss3 $\frac{\partial g(u)}{\partial x} = \frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial x}$ (chain rule)

ss4 $\frac{\partial f(g(u))}{\partial x} = \frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial x}$ (chain rule)

Properties of matrix derivatives cont.

Properties for scalar by scaler

ss1 $\frac{\partial(u+v)}{\partial x} = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial x}$

ss2 $\frac{\partial uv}{\partial x} = u\frac{\partial v}{\partial x} + v\frac{\partial u}{\partial x}$ (product rule)

ss3 $\frac{\partial g(u)}{\partial x} = \frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial x}$ (chain rule)

ss4 $\frac{\partial f(g(u))}{\partial x} = \frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial x}$ (chain rule)

Properties for vector by scaler

vs1 $\frac{\partial(a\mathbf{u}+\mathbf{v})}{\partial x} = a\frac{\partial \mathbf{u}}{\partial x} + \frac{\partial \mathbf{v}}{\partial x}$, where a is not a function of x.

vs2 $\frac{\partial Au}{\partial x} = A\frac{\partial u}{\partial x}$ where A is not a function of x.

vs3 $\frac{\partial \mathbf{u}^T}{\partial x} = \left(\frac{\partial \mathbf{u}}{\partial x}\right)^T$

vs4 $\frac{\partial g(\mathbf{u})}{\partial x} = \frac{\partial g(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial x}$ (chain rule)

vs5 $\frac{\partial f(g(\mathbf{u}))}{\partial x} = \frac{\partial f(g)}{\partial g}\frac{\partial g(\mathbf{u})}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial x}$ (chain rule)

Properties of matrix derivatives cont.

Properties for matrix by scaler

ms1 $\frac{\partial aU}{\partial x} = a\frac{\partial U}{\partial x}$, where a is not a function of x.

ms2 $\frac{\partial AUB}{\partial x} = A\frac{\partial U}{\partial x}B$ where A and B are not a function of x.

ms3 $\frac{\partial (U+V)}{\partial x} = \frac{\partial U}{\partial x} + \frac{\partial V}{\partial x}$

ms4 $\frac{\partial UV}{\partial x} = U\frac{\partial V}{\partial x} + \frac{\partial U}{\partial x}V$ (product rule)

Properties of matrix derivatives cont.

Properties for matrix by scaler

ms1 $\frac{\partial aU}{\partial x} = a\frac{\partial U}{\partial x}$, where a is not a function of x.

ms2 $\frac{\partial AUB}{\partial x} = A\frac{\partial U}{\partial x}B$ where A and B are not a function of x.

ms3 $\frac{\partial (U+V)}{\partial x} = \frac{\partial U}{\partial x} + \frac{\partial V}{\partial x}$

ms4 $\frac{\partial UV}{\partial x} = U\frac{\partial V}{\partial x} + \frac{\partial U}{\partial x}V$ (product rule)

Properties for scalar by vector

sv1 $\frac{\partial (au+v)}{\partial \mathbf{x}} = a\frac{\partial u}{\partial \mathbf{x}} + \frac{\partial v}{\partial \mathbf{x}}$, where a is not a function of x.

sv2 $\frac{\partial uv}{\partial \mathbf{x}} = u\frac{\partial v}{\partial \mathbf{x}} + v\frac{\partial u}{\partial \mathbf{x}}$ (product rule)

sv3 $\frac{\partial f(g(u))}{\partial \mathbf{x}} = \frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial \mathbf{x}}$ (chain rule)

sv4 $\frac{\partial \mathbf{u}^T\mathbf{v}}{\partial \mathbf{x}} = \mathbf{u}^T\frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^T\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$ (product rule)

sv5 $\frac{\partial \mathbf{u}^TA\mathbf{v}}{\partial \mathbf{x}} = \mathbf{u}^TA\frac{\partial \mathbf{v}}{\partial \mathbf{x}} + \mathbf{v}^TA^T\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$, where A is not a function of $\mathbf{x}$

Properties of matrix derivatives cont.

Properties for scalar by matrix

sm1 $\frac{\partial au}{\partial X} = a\frac{\partial u}{\partial X}$, where a is not a function of x.

sm2 $\frac{\partial(u+v)}{\partial X} = \frac{\partial u}{\partial X} + \frac{\partial v}{\partial X}$

sm3 $\frac{\partial uv}{\partial X} = u\frac{\partial v}{\partial X} + v\frac{\partial u}{\partial X}$ (product rule)

sm4 $\frac{\partial f(g(u))}{\partial X} = \frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial X}$ (chain rule)

Properties of matrix derivatives cont.

Properties for scalar by matrix

sm1 $\frac{\partial au}{\partial X} = a\frac{\partial u}{\partial X}$, where a is not a function of x.

sm2 $\frac{\partial (u+v)}{\partial X} = \frac{\partial u}{\partial X} + \frac{\partial v}{\partial X}$

sm3 $\frac{\partial uv}{\partial X} = u\frac{\partial v}{\partial X} + v\frac{\partial u}{\partial X}$ (product rule)

sm4 $\frac{\partial f(g(u))}{\partial X} = \frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial u}\frac{\partial u}{\partial X}$ (chain rule)

Properties for vector by vector

vv1 $\frac{\partial (a\mathbf{u}+\mathbf{v})}{\partial \mathbf{x}} = a\frac{\partial \mathbf{u}}{\partial \mathbf{x}} + \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$, where a is not a function of x.

vv2 $\frac{\partial A\mathbf{u}}{\partial \mathbf{x}} = A\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$, where A is not a function of $\mathbf{x}$.

vv3 $\frac{\partial f(g(\mathbf{u}))}{\partial \mathbf{x}} = \frac{\partial f(g)}{\partial g}\frac{\partial g(u)}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}$ (chain rule)

Zero-one matrices

Definition

A matrix all of whose entries are either 0 or 1 is called a zero-one matrix.

Zero-one matrices

Definition

A matrix all of whose entries are either 0 or 1 is called a zero-one matrix.

Operations

- $\wedge$: Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $m \times n$ zero-one matrices, $A \wedge B = [a_{ij} \wedge b_{ij}]$, i.e.,

$$(A \wedge B)_{ij} = \begin{cases} 1, & \text{if } a_{ij} = b_{ij} = 1; \\ 0, & \text{otherwise.} \end{cases}$$

- $\vee$: Let $A = [a_{ij}]$ and $B = [b_{ij}]$ be $m \times n$ zero-one matrices, $A \vee B = [a_{ij} \vee b_{ij}]$, i.e.,

$$(A \vee B)_{ij} = \begin{cases} 1, & \text{if } a_{ij} = 1 \text{ or } b_{ij} = 1; \\ 0, & \text{otherwise.} \end{cases}$$

Boolean product

Definition

If A is an $n \times m$ zero-one matrix and B is an $m \times p$ zero-one matrix, then $A \bigodot B = [c_{ij}] = \bigvee_{k=1}^{m}(a_{ik} \wedge b_{kj})$.

Boolean product

Definition

If A is an $n \times m$ zero-one matrix and B is an $m \times p$ zero-one matrix, then $A \bigodot B = [c_{ij}] = \bigvee_{k=1}^{m}(a_{ik} \wedge b_{kj})$.

Example

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

Solution

$A \bigodot B =$

$$\begin{pmatrix} (1 \wedge 1) \vee (0 \wedge 0) & (1 \wedge 1) \vee (0 \wedge 1) & (1 \wedge 0) \vee (0 \wedge 1) \\ (0 \wedge 1) \vee (1 \wedge 0) & (0 \wedge 1) \vee (1 \wedge 1) & (0 \wedge 0) \vee (1 \wedge 1) \\ (1 \wedge 1) \vee (0 \wedge 0) & (1 \wedge 1) \vee (0 \wedge 1) & (1 \wedge 0) \vee (0 \wedge 1) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Take-aways

Conclusions

- Sequence and summations
- Cardinality of set
- Matrix