

1. 희소행렬

행렬요소의 값이 대부분 0인 행렬을 희소행렬이라 한다.

다음은 희소행렬의 한 가지 예이다.

	1열	2열	3열	4열	5열	6열	7열	
1행	┌ 0	6	0	0	0	0	0	┐
2행	0	0	0	0	0	0	0	
3행	0	0	0	0	9	0	0	
4행	0	0	7	0	0	0	0	
5행	└ 0	0	0	0	0	8	0	┘

→ 0이 아닌 원소수 : 4개

- 주어진 행렬은 값이 대부분 0으로 구성되어 있다. 즉, 희소행렬이다.
- 희소행렬을 표현하는 방법은 여러 가지가 있을 수 있다.
- 그냥 단순하게 주어진 모습 그대로 2차원 배열로 표현할 수도 있다.

// 다음은 주어진 희소행렬을 있는 모습 그대로 2차원 배열로 표현한 것이다.(나쁜 방법)

0	6	0	0	0	0	0
0	0	0	0	0	0	0
0	0	0	0	9	0	0
0	0	7	0	0	0	0
0	0	0	0	0	8	0

필요한 기억공간 수 = 5 × 7 = 35

- 그런데, 희소행렬은 있는 모습 그대로인 2차원 배열구조가 아닌 다른 구조로 표현할 수 있다.
- 다른 구조는 0이 많은 것을 이용하여 또 다른 모습으로 표현하는 것이다.(희소행렬표현법)
- 즉, 행렬요소가 0이 아닌 것을 이용한 표현법을 희소행렬표현법이라 한다.(좋은 방법)

// 희소행렬(sparse matrix)과 그 표현법

행렬요소 대부분이 0인 행렬을 희소행렬이라 한다. 다음은 희소행렬의 한 가지 예이다.

〈희소행렬〉

	1열	2열	3열	4열	5열	6열	7열
1행	0	6	0	0	0	0	0
2행	0	0	0	0	0	0	0
3행	0	0	0	0	9	0	0
4행	0	0	7	0	0	0	0
5행	0	0	0	0	0	8	0

→ 0이 아닌 원소수 : 4개

↓
↓ 2차원 배열구조의 희소행렬표현법으로 나타내면
↓

〈희소행렬표현법〉

5	7	4	→ 행렬이 5행 7열 구조이며, 0이 아닌 요소가 4개라는 뜻이다
1	2	6	→ 1행 2열의 행렬요소 값이 6이라는 뜻이다
3	5	9	→ 3행 5열의 행렬요소 값이 9라는 뜻이다
4	3	7	→ 4행 3열의 행렬요소 값이 7라는 뜻이다
5	6	8	→ 5행 6열의 행렬요소 값이 8라는 뜻이다

• 필요한 기억공간 수 = $3 \times 5 = 15$

- 첫 번째 행에는 행렬구조와 0이 아닌 행렬요소의 개수를 기억시킨다.
- 두 번째 행부터는 (행위치, 열위치, 값)을 표현한다.
- 희소행렬표현법은 행과 열의 위치로 행렬요소를 나타낼 수 있다는 것을 기본원칙으로 한다.
- 희소행렬표현법은 행렬요소가 0이 아닌 것을 이용한 표현법이다.
- 희소행렬 표현의 가장 큰 목적은 기억장소를 절약하기 위한 것이다.
- 즉, 15개의 기억공간으로 5행 7열의 희소행렬을 표현하였다.
- 희소행렬표현법은 배열 또는 연결구조로 표현 가능하다.
- 현재, 희소행렬표현법은 배열로 표현된 것이 출제되고 있다.

기출문제 분석

1. 값이 0인 원소들의 비율이 90%인 2,000×2,000 희소행렬 표현에 관한 설명 중 옳은 것은?
[2007년 국가 7급]

- ① 2차원 배열로 모든 원소들을 표현하는 것이 저장 관점에서 효율적이다.
- ② 2차원 배열로 0이 아닌 원소들만 표현하기 위해서 2,000개의 저장공간이 필요하다.
- ③ 0이 아닌 원소들만 연결리스트로 표현하는 것이 배열로 표현하는 것보다 전치행렬 연산에 효율적이다.
- ④ 0이 아닌 원소들만 3원소쌍(행의 인덱스, 열의 인덱스, 값)으로 배열에 저장하는 것이 저장 관점에서 효율적이다.

☞ 희소행렬 표현

- 0이 아닌 원소 수 = 2,000 × 2,000 × 0.1 = 400,000
- 0이 아닌 원소를 표현하기 위해서 구조체 배열로 400,001개의 저장공간이 필요하다.

```
struct matrix
{
    int row;      → 행의 인덱스
    int col;      → 열의 인덱스
    int value;    → 0이 아닌 원소의 값
} a[400001];     → 3원소쌍의 구조체 배열구조
```

◆ 전치행렬

1	2	3
4	5	6
7	8	9

↓ 전치행렬 : 행은 열로, 열은 행으로

1	4	7
2	5	8
3	6	9

- 전치행렬 연산은 2차원 배열로 표현하여 처리하면 효율적으로 구현할 수 있다.

2. 다음 5×5 희소행렬(sparse matrix)을 구조체 배열로 표현하기 위해 <규칙>의 표현 규칙을 사용하여 <코드>의 C 언어 코드를 작성하였다. <코드>의 ㉠에 들어갈 내용은? (단, 행렬의 행 번호와 열 번호는 0부터 시작한다) [2018년 국가 7급]

0	5	2	0	0
0	0	0	1	0
0	0	0	0	0
3	0	0	0	0
0	0	0	0	0

```
-----<코 드>-----
typedef struct {
    int row;
    int col;
    int val;
} term;
term A[ ] = _____ ㉠ _____ ;
-----<규 칙>-----
```

- 첫 번째 구조체인 A[0]의 row, col, val에는 희소행렬의 행 개수, 열 개수, 값이 0이 아닌 원소 개수를 순서대로 저장한다.
- 두 번째부터의 나머지 구조체는 값이 0이 아닌 희소행렬 원소를 표현하기 위해 사용되며, 각 구조체의 row, col, val에 해당 원소의 행 번호, 열 번호, 값을 순서대로 저장한다.
- 값이 0이 아닌 희소행렬 원소 정보는 행 번호의 오름차순으로 구조체 배열 A에 저장되며, 행 번호가 같은 경우에는 열 번호의 오름차순으로 저장된다.

- ① {{4, 4, 4}, {0, 1, 5}, {0, 2, 2}, {1, 3, 1}, {3, 0, 3}}
- ② {{4, 4, 4}, {3, 0, 3}, {0, 1, 5}, {0, 2, 2}, {1, 3, 1}}
- ③ {{5, 5, 4}, {0, 1, 5}, {0, 2, 2}, {1, 3, 1}, {3, 0, 3}}
- ④ {{5, 5, 4}, {3, 0, 3}, {0, 1, 5}, {0, 2, 2}, {1, 3, 1}}

♣ 5×5 희소행렬

㉠에 들어갈 내용은 희소행렬을 희소행렬표현으로 나타낸 것을 초기화하는 것이다.

0	5	2	0	0
0	0	0	1	0
0	0	0	0	0
3	0	0	0	0
0	0	0	0	0

----->
희소행렬표현

5	5	4
0	1	5
0	2	2
1	3	1
3	0	3

• 문제만 복잡함

3. X, Y, Z 는 $m \times n$ 희소행렬에서 원소의 값이 0이 아닌 L개의 각 원소를 표현하기 위해 <행, 열, 값> 3원소 쌍을 사용하는 2차원 배열이며, 변환규칙은 (가)와 같다. 행렬 P와 Q가 (나)와 같이 X, Y로 각각 표현되었을 때, P와 Q를 곱한 행렬 R을 Z로 표현할 경우, 이에 대한 설명으로 옳지 않은 것은? (단, 행렬의 행 번호와 열 번호는 0부터 시작한다) [2014년 국가 7급]

(가)-----

- 행우선 배열이며, 크기는 $(L+1) \times 3$ 이다.
- 첫 번째 행의 3개 열에는 희소행렬에 대한 행의 크기(m), 열의 크기(n), 0이 아닌 원소의 수(L)를 순서대로 저장한다.
- 나머지 L개의 각 행에는 0이 아닌 L개의 원소들에 대한 정보가 저장되며, 해당 행의 각 열에는 원소들의 행 번호, 열 번호, 저장된 값이 각각 저장된다.
- 원소들에 대한 정보를 저장할 때는 원소들의 행 번호와 열 번호가 차례대로 오름차순이 되도록 저장한다.

(나)-----

X	행\열	0	1	2
	0	3	4	3
	1	0	3	4
	2	1	1	7
	3	2	3	1

Y	행\열	0	1	2
	0	4	2	3
	1	0	1	1
	2	3	0	3
	3	3	1	2

- ① $Z[0][1] = 2$ 이다.
- ② $Z[0][2] = 4$ 이다.
- ③ $Z[1][0] = 0$ 이다.
- ④ $Z[3][2] = 2$ 이다.

☞ 희소행렬 - 행렬의 곱셈

• 주어진 문제는 행렬의 곱셈을 알고 있어야 한다.(이런 유형으로는 처음 출제되었다)

P	0	0	0	4
	0	7	0	0
	0	0	0	1
	×			

Q	0	1
	0	0
	0	0
	3	2

↓ $R = P \cdot Q$ (행렬의 곱셈)

R	$0 \cdot 0 + 0 \cdot 0 + 0 \cdot 0 + 4 \cdot 3 = 12$	$0 \cdot 1 + 0 \cdot 0 + 0 \cdot 0 + 4 \cdot 2 = 8$
	$0 \cdot 0 + 7 \cdot 0 + 0 \cdot 0 + 0 \cdot 3 = 0$	$0 \cdot 1 + 7 \cdot 0 + 0 \cdot 0 + 0 \cdot 2 = 0$
	$0 \cdot 0 + 0 \cdot 0 + 0 \cdot 0 + 1 \cdot 3 = 3$	$0 \cdot 1 + 0 \cdot 0 + 0 \cdot 0 + 1 \cdot 2 = 2$

↓ R을 희소행렬 Z으로 표현하면

6 <http://cafe.daum.net/pass365>(홍재연)

Z	행\열	0	1	2
	0	3	2	4
	1	0	0	12
	2	0	1	8
	3	2	0	3
	4	2	1	2

- $Z[3][2] = 2$ 이다. (\times) $\rightarrow Z[3][2] = 3$ 이다.
- 이런 문제는 1분 안에 풀 수가 없다. 내용을 잘 알 때, 과연 몇 분 만에 풀 수 있을까?

정답 : ④

4. 희소행렬에 대한 설명으로 옳지 않은 것은? [2011년 국가 7급]

- ① 대부분의 원소 값이 0으로 구성되어 있다.
- ② 2차원 배열로 표현하면 특정 항목의 접근이 용이하다.
- ③ 연결리스트 구조로 표현하더라도 행렬의 덧셈 연산을 할 수 있다.
- ④ 연결리스트 구조로 표현하면 기억공간을 낭비하게 된다.

☞ 희소행렬 표현법

- 희소행렬은 배열 또는 연결리스트 구조로 표현할 수 있다.
- 희소행렬 표현법은 단지, 기억공간을 절약하기 위한 표현법이다.

정답 : ④

5. $m * n$ 크기의 정수 값 희소행렬을 정수형 배열에 저장하고자 한다. 가장 효과적인 저장 방법을 사용할 때, 필요한 배열의 크기는?(단, t 는 0이 아닌 희소행렬 원소의 개수이며, 배열의 크기는 배열원소의 수를 의미한다) [2010년 국가 7급]

- ① $m * n$
- ② t
- ③ $m + n - 1$
- ④ $3t$

☞ 희소행렬표현법 - (행위치, 열위치, 0이 아닌 원소 값)

- $3t+3 \rightarrow t$ 는 0이 아닌 원소 수, $+3$ 은 행렬구조를 표현하기 위한 기억공간 수
- 그런데, $3t+3$ 이 없다. 정답은 그냥 근사치인 $3t$ 로 하는 수밖에

정답 : ④