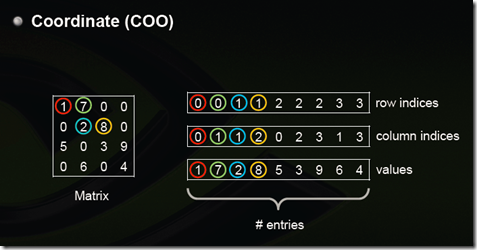
稀疏矩阵的存储格式（Sparse Matrix Storage Formats）

备注：行和列都是0索引开始。

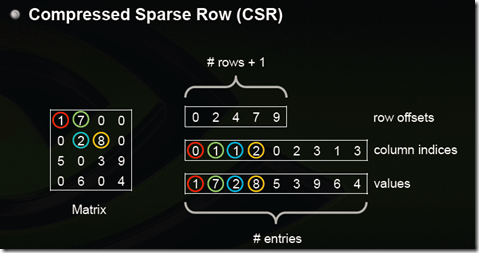
稀疏矩阵是指矩阵中的元素大部分是0的矩阵，事实上，实际问题中大规模矩阵基本上都是稀疏矩阵，很多稀疏度在90%甚至99%以上。因此我们需要有高效的稀疏矩阵存储格式。

（1）**ST** 格式（Sparse Triplet Format）或Coordinate（COO）

[](https://images0.cnblogs.com/blog/354318/201502/042300488598079.png)

这是最简单的一种格式，每一个元素需要用一个三元组来表示，分别是（行号，列号，数值），对应上图右边的一列。这种方式简单，但是记录单信息多（行列），每个三元组自己可以定位，因此空间不是最优。

（2）Compressed Sparse Row (CSR)

[](https://images0.cnblogs.com/blog/354318/201502/042300502345067.png)

CSR是比较标准的一种，也需要三类数据来表达：数值，列号，以及行偏移。CSR不是三元组，而是整体的编码方式。数值和列号与COO一致，表示一个元素以及其列号，行偏移表示某一行的第一个元素在values里面的起始偏移位置。如上图中，第一行元素1是0偏移，第二行元素2是2偏移，第三行元素5是4偏移，第4行元素6是7偏移。在行偏移的最后补上矩阵总的元素个数，本例中是9。

CSC是和CSR相对应的一种方式，即按列压缩的意思。

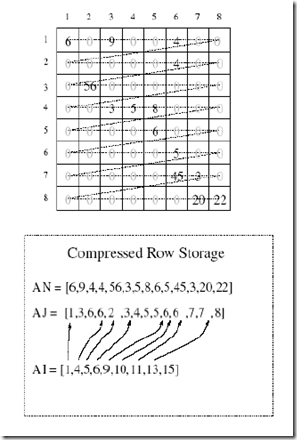
以上图中矩阵为例：

Values：        [1 5 7 2 6 8 3 9 4]

Row Indices：[0 2 0 1 3 1 2 2 3]

Column Offsets：[0 2 5 7 9]

 再来看一个CSR的例子：

[](https://images0.cnblogs.com/blog/354318/201502/042300515627039.png)

测试结果：

Sparse Triplet (ST) header information:

Minimum row index I\_MIN = 1

Maximum row index I\_MAX = 5

Minimum col index J\_MIN = 1

Maximum col index J\_MAX = 5

Number of rows M = 5

Number of columns N = 5

Number of nonzeros NST = 12

The matrix in ST format:

# I J A

---- ---- ---- --------------

0 0 0 2

1 0 1 3

2 1 0 3

3 1 2 4

4 1 4 6

5 2 1 -1

6 2 2 -3

7 2 3 2

8 3 2 1

9 4 1 4

10 4 2 2

11 4 4 1

Number of CC values = 12

CC Matrix:

# I J A

---- ---- ---- ----------------

0 0 0 2

1 1 0 3

2 0 0 3

3 2 0 -1

4 4 0 4

5 1 1 4

6 2 1 -3

7 3 1 1

8 4 1 2

9 2 2 2

10 1 3 6

11 4 3 1

TEST02

Convert a sparse matrix from ST to CC format.

ST: sparse triplet, I, J, A.

CC: compressed column, I, CC, A.

This matrix is read from the file 'west\_st.txt'

Sparse Triplet (ST) header information:

Minimum row index I\_MIN = 1

Maximum row index I\_MAX = 67

Minimum col index J\_MIN = 1

Maximum col index J\_MAX = 67

Number of rows M = 67

Number of columns N = 67

Number of nonzeros NST = 299

Number of CC values = 294

TEST03

Convert a sparse matrix from ST to CC format.

ST: sparse triplet, I, J, A.

CC: compressed column, I, CC, A.

The ST matrix is read from the file 'west\_st.txt'

and the CC matrix is written to the files:

'west\_icc.txt',

'west\_ccc.txt', and

'west\_acc.txt'.

Sparse Triplet (ST) header information:

Minimum row index I\_MIN = 1

Maximum row index I\_MAX = 67

Minimum col index J\_MIN = 1

Maximum col index J\_MAX = 67

Number of rows M = 67

Number of columns N = 67

Number of nonzeros NST = 299

Number of CC values = 294