

14<sup>th</sup> february

A = [ 16 elements of single precision numbers]

A = 4x4 and B = 4x1 (vector)

A\*B = 4x1

Diff vector – storing the number in the matrix which are referenced by pointers. The pointers have spatial locality but not the data.

Sparse matrix

The matrix in which most of the elements in the matrix are zero

sparsity = Number of zero elements / Number of elements

Used in the graphs, convolution network.

Two approaches to determine the sparse matrix zero elements

1.) efficient updates

2.) read-only access

1.) Efficient updates

Dictionary of keys – Map the row-column pair to the non-zero element

Compressed sparse row format – to store the column and row of the dictionary keys we have mapped.

IA[]- length of the # of the non-zero

JA[] - column index for non zero

When do I go sparse?

It is the function of density = #of non zero < (m(n-1) -1)/2

C= A B

for(i=0; i < n ; i++)

    for(j=0; j < n; j++)

        for(k =0; k < n ; k++)

            C[i][j] +=A[i][k] \* B[k][j]

B how many reads of an elements =  $n^3$

A how many  $n^2$

C read/written form main memory  $2n^2$