Storage Analysis

The purpose of this document is to hold the derived storage requirements of different matrix storage formats (COO, CSR, etc.). Equations derived here can then be used to inform plots for the presentation itself.

Dense Format

Requires a single 2 dimensional array to hold all matrix entries, zero and non-zero.

For an $n \times n$ matrix with $0 \leq m \leq n^2$ non-zero elements, dense format requires

n^2 elements of the type used in the matrix

COO Format

Requires 3 indices (row, col, val) split across 3 arrays per non-zero matrix element.

For an $n \times n$ matrix with $0 \leq m \leq n^2$ non-zero elements, COO format requires

$2m$ integers + m elements of the type used in the matrix

row + col ↑ ↑ values

If the number of non-zero elements is expressed as a density decimal $0 \leq d \leq 1$ (where $d = 0$ implies all zero entries and $d = 1$ implies no zero entries) instead of as the number of non-zero, the number of elements required to store the matrix is as follows:

$2d \cdot n^2$ integers + $d \cdot n^2$ elements of the type used by the matrix

row + col ↑ ↑ values

To access the Google Docs version of this file, go to the link below:
https://docs.google.com/document/d/1qfStOPBqG2O3u0i8hAnnkBAOTh6TvtSIB_6fu7A_M64/edit?usp=sharing

CSR Format

Requires 1 value and column index per non-zero matrix element. Number of row pointer values equals number of rows + 1 (determined experimentally, is there an analytical reason why?).

For an $n \times n$ matrix with $0 \leq m \leq n^2$ non-zero elements, CSR format requires

$(n + 1) + m$ integers + m elements of the type used in the matrix

row ↑ col ↑ ↑ values

If the number of non-zero elements is expressed as a density decimal $0 \leq d \leq 1$ (where $d = 0$ implies all zero entries and $d = 1$ implies no zero entries) instead of as the number of non-zero, the number of elements required to store the matrix is as follows:

$(n + 1) + d \cdot n^2$ integers + $d \cdot n^2$ elements of the type used by the matrix

row ↑ col ↑ ↑ values

CSC Format

Requires 1 value and row index per non-zero matrix element. Number of column pointer values equals number of columns + 1 (determined experimentally, is there an analytical reason why?).

For an $n \times n$ matrix with $0 \leq m \leq n^2$ non-zero elements, CSC format requires

$(n + 1) + m$ integers + m elements of the type used in the matrix

col ↑ row ↑ ↑ values

If the number of non-zero elements is expressed as a density decimal $0 \leq d \leq 1$ (where $d = 0$ implies all zero entries and $d = 1$ implies no zero entries) instead of as the number of non-zero, the number of elements required to store the matrix is as follows:

$(n + 1) + d \cdot n^2$ integers + $d \cdot n^2$ elements of the type used by the matrix

col ↑ row ↑ ↑ values

To access the Google Docs version of this file, go to the link below:
https://docs.google.com/document/d/1qfStOPBqG2O3u0i8hAnnkBAOTh6TvtSIB_6fu7A_M64/edit?usp=sharing