

Komprimierung von dünnbesetzten Matrizen

Raffael Lorup, Paolo Di Stolfo

Präsentation, am 21.01.2011

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Definitionen

Motivation und Ziel

Motivation und Ziel

Komprimierung

Komprimierung

Matrix-Vektor-Produkt

Matrix-Vektor-
Produkt

Vergleich der Speicherformate

Vergleich der
Speicherformate

Dünnbesetzte Matrix (*sparse matrix*)

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definition: Dünnbesetzte Matrix

Eine $m \times n$ -Matrix \mathbf{A} heißt *dünnbesetzt*, wenn gilt:

- ▶ *wenige* Elemente in \mathbf{A} ungleich 0
- ▶ auf \mathbf{A} sind spezielle Techniken anwendbar, die die große Anzahl von 0 ausnützen

Definitionen

Motivation und Ziel

Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Definition

- ▶ Nz heißt Anzahl der Elemente in \mathbf{A} , die ungleich 0 sind

1. Speicherung: Platzersparnis

- ▶ Graphen mit wenigen Kanten: Adjazenzmatrizen
- ▶ mathematische Modelle liefern dünnbesetzte Gleichungssysteme

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

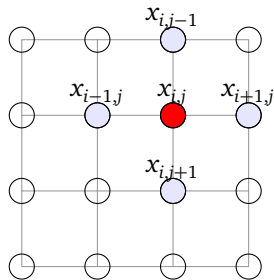
Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

1. Speicherung: Platzersparnis

- ▶ Graphen mit wenigen Kanten: Adjazenzmatrizen
- ▶ mathematische Modelle liefern dünnbesetzte Gleichungssysteme
z.B. Temperatur auf Platte



$$x_{i,j} = \frac{1}{4} (x_{i,j-1} + x_{i-1,j} + x_{i,j+1} + x_{i+1,j})$$

1. Speicherung: Platzersparnis

- ▶ Graphen mit wenigen Kanten: Adjazenzmatrizen
- ▶ mathematische Modelle liefern dünnbesetzte Gleichungssysteme
z.B. Temperatur auf Platte
- ▶ $n := 10^7$ Gleichungen = Gitterpunkte
 $\Rightarrow 10^{14}$ Koeffizienten
- ▶ naive Datenstruktur braucht dann 400 TB
(für Gleitkommazahlen)

1. Speicherung: Platzersparnis

- ▶ Graphen mit wenigen Kanten: Adjazenzmatrizen
- ▶ mathematische Modelle liefern dünnbesetzte Gleichungssysteme
z.B. Temperatur auf Platte
- ▶ $n := 10^7$ Gleichungen = Gitterpunkte
 $\Rightarrow 10^{14}$ Koeffizienten
- ▶ naive Datenstruktur braucht dann 400 TB
(für Gleitkommazahlen)

2. Operationen: Performance-Gewinn

- ▶ Matrix-Vektor-Produkt $y := Ax$
- ▶ naiver Gauß: braucht $\frac{2}{3}n^3$ Operationen

Überlegungen zur Wahl des Speicherformats

- ▶ Speicherverbrauch
- ▶ Ausführungsgeschwindigkeit der Operationen
- ▶ anzuwendender Algorithmus
- ▶ Struktur der Matrix
- ▶ Rechnerarchitektur
- ▶ ...

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

- ▶ Nichtnull-Elemente festhalten
 - ▶ geeignete Reihenfolge
 - ▶ zugehörige Position

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

- ▶ Nichtnull-Elemente festhalten
 - ▶ geeignete Reihenfolge
 - ▶ zugehörige Position
- ▶ Allgemeine Formate:
 - ▶ Koordinaten-Format (COO)
 - ▶ Compressed Row Storage (CRS)
 - ▶ Compressed Column Storage (CCS)
 - ▶ Modified Row Storage (MRS)
- ▶ Spezielle Formate:
 - ▶ Compressed Diagonal Storage (CDS)
 - ▶ Skyline Storage (SKS)

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format
Compressed Row Storage
Compressed Column Storage
Modified Row Storage
Compressed Diagonal Storage
Skyline Storage

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Koordinaten-Format (COO)

- Vorteil: Einfachheit

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Koordinaten-Format (COO)

- ▶ Vorteil: Einfachheit
- ▶ Speicheraufwand $3Nz$

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0

Values

11	12	21	22	33	34	43	44	54	55	65
----	----	----	----	----	----	----	----	----	----	----

RowIndices

1	1	2	2	3	3	4	4	5	5	6
---	---	---	---	---	---	---	---	---	---	---

ColumnIndices

1	2	1	2	3	4	3	4	4	5	5
---	---	---	---	---	---	---	---	---	---	---

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Compressed Row Storage (CRS)

- ▶ zeilenweiser Durchgang

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

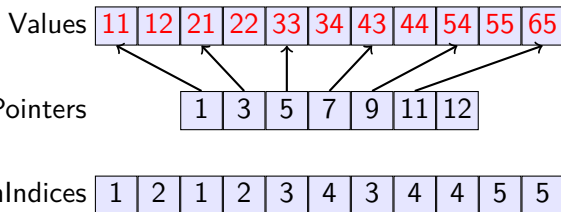
Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Compressed Row Storage (CRS)

- ▶ zeilenweiser Durchgang
- ▶ Zeile von $\text{RowPointers}[i]$ bis $\text{RowPointers}[i + 1] - 1$
- ▶ Speicheraufwand ist $2Nz + (m + 1)$

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0



Komprimierung von dünnbesetzten Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-Produkt

Vergleich der Speicherformate

Compressed Column Storage (CCS)

► wie *CRS* auf A^T

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

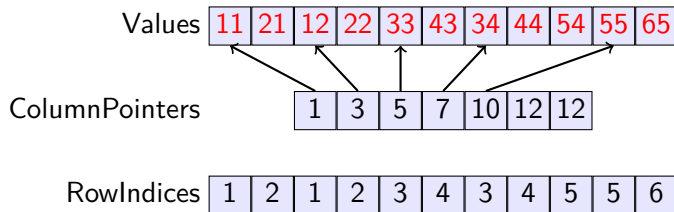
Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Compressed Column Storage (CCS)

- ▶ wie *CRS* auf A^T
- ▶ Speicherbedarf ist $2Nz + (n + 1)$

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0



Komprimierung von dünnbesetzten Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-Produkt

Vergleich der Speicherformate

Modified Row Storage (*MRS*)

- ▶ A sei $n \times n$ -Matrix
- ▶ Hauptdiagonale
- ▶ zeilenweise

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Modified Row Storage (MRS)

- ▶ A sei $n \times n$ -Matrix
- ▶ Hauptdiagonale
- ▶ zeilenweise
- ▶ Speicheraufwand ist $2(Nz + 1 + \#Nullen \text{ in Diagonale})$

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0

Komprimierung von dünnbesetzten Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

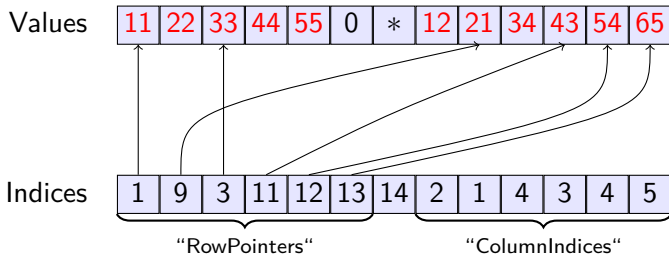
Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-Produkt

Vergleich der Speicherformate



Compressed Diagonal Storage (*CDS*)

- ▶ A sei $n \times n$ -Matrix
- ▶ konstante Bandbreite

Bandbreite	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Compressed Diagonal Storage (CDS)

- ▶ A sei $n \times n$ -Matrix
- ▶ konstante Bandbreite
- ▶ Speicherverbrauch liegt bei

$n \times \text{Bandbreite}$

Bandbreite	1	2	3	4	5	6	
1	11	12	0	0	0	0	
2	21	22	0	0	0	0	
3	0	0	33	34	0	0	
4	0	0	43	44	0	0	
5	0	0	0	54	55	0	
6	0	0	0	0	65	0	1
					-1	0	

Values	*	21	0	43	54	65	-1	Diagonals
	11	22	33	44	55	0	0	
	12	0	34	0	0	*	1	

Komprimierung von dünnbesetzten Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-Produkt

Vergleich der Speicherformate

- ▶ spezielle Skyline-Struktur

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des
Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

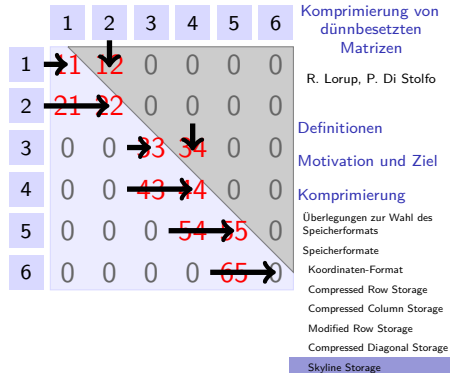
Skyline Storage

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Skyline Storage

- ▶ spezielle Skyline-Struktur



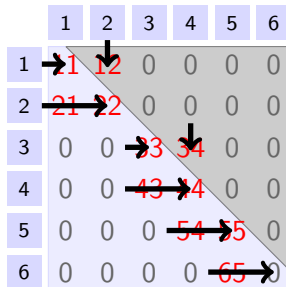
Matrix-Vektor-Produkt

Vergleich der Speicherformate

Skyline Storage

- ▶ spezielle Skyline-Struktur
- ▶ Speicherverbrauch ist

$$Nz + 2(n + 1) + \#Skyline\text{-Nullen}$$



Komprimierung von dünnbesetzten Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-Produkt

Vergleich der Speicherformate

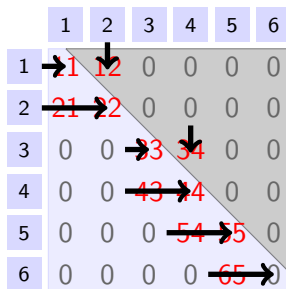
LowerValues 11 21 22 33 43 44 54 55 65 0

LowerRowPointers 1 2 4 5 7 9 11

Skyline Storage

- ▶ spezielle Skyline-Struktur
- ▶ Speicherverbrauch ist

$$Nz + 2(n + 1) + \#Skyline-Nullen$$



Komprimierung von dünnbesetzten Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Überlegungen zur Wahl des Speicherformats

Speicherformate

Koordinaten-Format

Compressed Row Storage

Compressed Column Storage

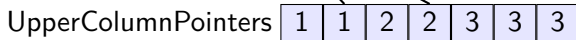
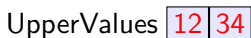
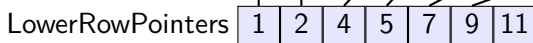
Modified Row Storage

Compressed Diagonal Storage

Skyline Storage

Matrix-Vektor-Produkt

Vergleich der Speicherformate



Musteroperation: Matrix-Vektor-Produkt

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

für $m \times n$ -Matrix \mathbf{A} und Vektor $x \in \mathbb{K}^n$:

$$y := Ax = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$
$$= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}$$

Definitionen

Motivation und Ziel

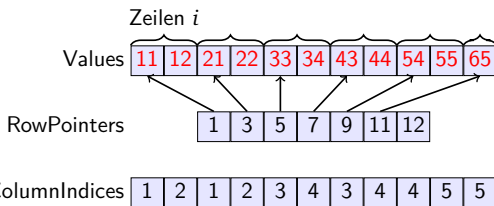
Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

CRS - $y := Ax$

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0



Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

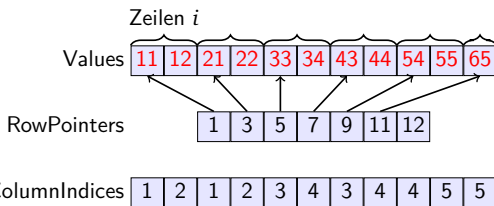
Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

CRS - $y := Ax$

	1	2	3	4	5	6
1	11	12	0	0	0	0
2	21	22	0	0	0	0
3	0	0	33	34	0	0
4	0	0	43	44	0	0
5	0	0	0	54	55	0
6	0	0	0	0	65	0



```
for  $i \leftarrow 1$  to #RowPointers - 1 do  
  for  $j \leftarrow$  RowPointers[ $i$ ] to RowPointers[ $i + 1$ ] - 1 do  
     $y[i] \leftarrow y[i] +$  Values[ $j$ ]  $\cdot x$ [ColumnIndices[ $j$ ]]  
  end for  
end for
```

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

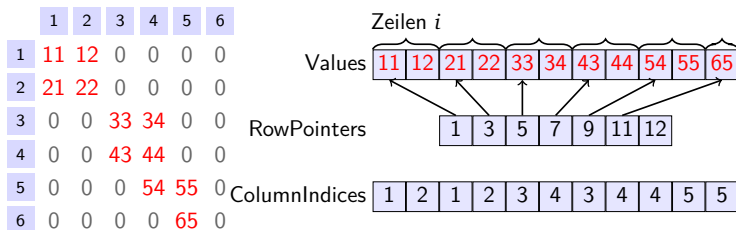
Motivation und Ziel

Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

CRS - $y := Ax$



```
for  $i \leftarrow 1$  to #RowPointers - 1 do
  for  $j \leftarrow$  RowPointers[ $i$ ] to RowPointers[ $i + 1$ ] - 1 do
     $y[i] \leftarrow y[i] + \text{Values}[j] \cdot x[\text{ColumnIndices}[j]]$ 
  end for
end for
```

- *Kosten:*
($2Nz$) Fließkomma-Operationen
vs. vorher ca. ($2mn$)

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Beispielhafter Vergleich der Speicherformate

Durchschnittlicher Speicherverbrauch je Form verglichen mit der ursprünglichen Matrix¹

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

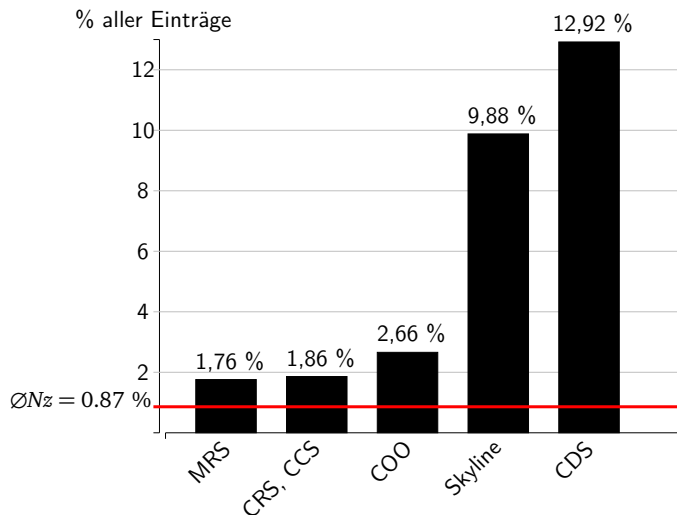
Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

¹nach [Überhuber], S. 405f.

Beispielhafter Vergleich der Speicherformate

Durchschnittlicher Speicherverbrauch je Form verglichen mit der ursprünglichen Matrix¹



Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

¹nach [Überhuber], S. 405f.

Vielen Dank für die Aufmerksamkeit!

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Literaturverzeichnis I



Aydin, B., T., F. J., Matteo, F., R., G. J., and E., L. C.

Parallel sparse matrix-vector and matrix-transpose-vector multiplication using compressed sparse blocks.

In *Proceedings of the twenty-first annual symposium on Parallelism in algorithms and architectures* (New York, NY, USA, 2009), SPAA '09, ACM, pp. 233–244.



Bai, Z., Demmel, J., Dongarra, J., Ruhe, A., and van der Vorst, H., Eds.

Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide.
SIAM, Philadelphia, 2000.



Barrett, R., Berry, M., Chan, T. F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., and van der Vorst, H.

Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd ed.
SIAM, Philadelphia, PA, 1994.



Eijkhout, V.

Distributed Sparse Data Structures for Linear Algebra Operations.
Department of Computer Science, University of Tennessee, Knoxville, 1992.



Gundersen, G., and Steihaug, T.

Data structures in Java for matrix computations: Research articles.
Concurr. Comput. : Pract. Exper. 16 (July 2004), 799–815.



I.S. Duff, A.M. Erisman, J. R.

Direct Methods for Sparse Matrices.
Oxford University Press, New York, 1986.
ISBN: 0-198-53408-6.



Montagne, E., and Ekambaram, A.

An optimal storage format for sparse matrices.
Inf. Process. Lett. 90 (April 2004), 87–92.

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Literaturverzeichnis II



Pooch, U. W., and Nieder, A.

A survey of indexing techniques for sparse matrices.
ACM Comput. Surv. 5 (June 1973), 109–133.



Saad, Y.

Iterative Methods for Sparse Linear Systems, second ed.
Society for Industrial and Applied Mathematics, 1996.
ISBN-13: 978-0-898715-34-7.



Silva, M.

Sparse matrix storage revisited.
In *Proceedings of the 2nd conference on Computing frontiers* (New York, NY, USA, 2005), CF '05, ACM, pp. 230–235.



Smailbegovic, F., Gaydadjiev, G. N., and Vassiliadis, S.

Sparse matrix storage format.
In *Proceedings of the 16th Annual Workshop on Circuits, Systems and Signal Processing, ProRisc 2005* (November 2005), pp. 445–448.



Stathis, P. T.

Sparse Matrix Vector Processing Formats.
TU Delft, Delft University of Technology (NL), 2004.
Doctoral thesis.



Überhuber, C.

Computer-Numerik 2.
Springer-Verlag Berlin Heidelberg, 1995.
ISBN 3-540-59152-4.

Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

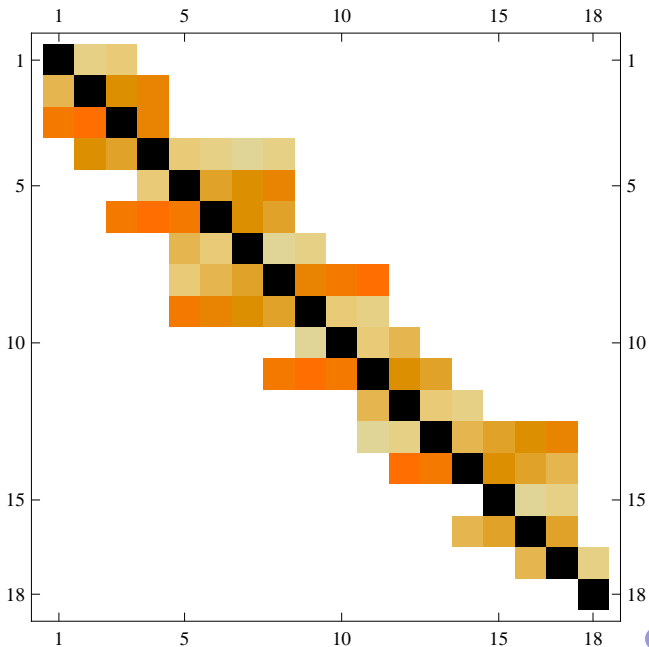
Motivation und Ziel

Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

Beispiel: Skyline-Matrix



Komprimierung von
dünnbesetzten
Matrizen

R. Lorup, P. Di Stolfo

Definitionen

Motivation und Ziel

Komprimierung

Matrix-Vektor-
Produkt

Vergleich der
Speicherformate

► Zurück