Jacobian Compression II

Graph Coloring

Uwe Naumann

Informatik 12:
Software and Tools for Computational Engineering (STCE)

RWTH Aachen University

Contents

Outline

Objective

- Introduction to direct Jacobian compression based on graph coloring.

Learning Outcomes

- You will understand
  - formulation of Jacobian compression problems as graph coloring problems
  - heuristics for distance-1 graph coloring
- You will be able to
  - derive graph representations of matrices
  - apply heuristics for distance-1 graph coloring

Outline

For a given sparsity pattern

$$P = P(F') \in \{0,1\}^{m \times n} \quad \text{of} \quad F' = F(\mathbf{x}) \in \boldsymbol{R}^{m \times n}$$

find

$$X^{(1)} \in \{0,1\}^{m \times n^{(1)}}$$

with minimal $n^{(1)}$ such that F' can be recovered by direct substitution from

$$\boldsymbol{R}^{m \times n^{(1)}} \ni Y^{(1)} = F' \cdot X^{(1)} \ .$$

Example:

$$\begin{pmatrix} a? & b? & \\ & c? & d? \\ e? & & \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & b \\ d & c \\ e & \end{pmatrix}$$

For a given sparsity pattern

$$P = P(F') \in \{0,1\}^{m \times n} \quad \text{of} \quad F' = F(\mathbf{x}) \in \boldsymbol{R}^{m \times n}$$

find

$$Y_{(1)} \in \{0,1\}^{m_{(1)} \times n}$$

with minimal $m_{(1)}$ such that F' can be recovered by direct substitution from

$$\boldsymbol{R}^{m_{(1)} \times n} \ni X_{(1)} = Y_{(1)} \cdot F' \; .$$

Example:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a? & b? & \\ & c? & d? \\ e? & & \end{pmatrix} = \begin{pmatrix} a & b \\ d & c \\ e & \end{pmatrix}$$

Find

$$X^{(1)} \in \{0,1\}^{m \times n^{(1)}} \quad \text{and} \quad Y_{(1)} \in \{0,1\}^{m_{(1)} \times n}$$

with minimal $n^{(1)} + \mathcal{R} \cdot m_{(1)}$, where $\mathcal{R}$ denotes the cost of an adjoint relative to a tangent, such that F' can be recovered by direct substitution from

$$\boldsymbol{R}^{m \times n^{(1)}} \ni Y^{(1)} = F' \cdot X^{(1)} \quad \text{and} \quad \boldsymbol{R}^{m_{(1)} \times n} \ni X_{(1)} = Y_{(1)} \cdot F' \ .$$

Example:

$$\begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} a? & b? & c? \\ d & e & \\ f & & g \end{pmatrix} = \begin{pmatrix} a & b & c \end{pmatrix}$$

$$\begin{pmatrix} a & b & c \\ d? & e? & \\ f? & & g \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b+c \\ d & e \\ f & g \end{pmatrix}$$

Outline

Software and Tools for Computational Engineering — RWTH AACHEN UNIVERSITY

- column compression problem → column intersection or bipartite graphs

$$\boldsymbol{R}^{m\times n^{(1)}} \ni Y^{(1)} = F' \cdot X^{(1)} \;.$$

- row compression problem → row intersection or bipartite graphs

$$\boldsymbol{R}^{m_{(1)}\times n} \ni X_{(1)} = Y_{(1)} \cdot F' \;.$$

- bi-directional compression problem → bipartite graph

$$\boldsymbol{R}^{m\times n^{(1)}} \ni Y^{(1)} = F' \cdot X^{(1)} \quad \text{and} \quad \boldsymbol{R}^{m_{(1)}\times n} \ni X_{(1)} = Y_{(1)} \cdot F' \;.$$

- symmetric compression problem (e.g. Hessian F'' →) adjacency graph

$$\boldsymbol{R}^{n\times n^{(2)}} \ni X_{(1)}^{(2)} = \underbrace{y_{(1)}}_{:=1} \cdot F'' \cdot X^{(2)} \;.$$

The undirected column intersection graph of a matrix $A = (a_{i,j}) \in \mathbb{R}^{m \times n}$ is defined as $G_c(A) = (V, E)$ where $V = \{0, \ldots, n-1\}$ and $(i,j) \in E \Leftrightarrow \exists k : a_{k,i} \neq 0 \wedge a_{k,j} \neq 0$.

The undirected row intersection graph of A is defined as $G_r(A) = (V, E)$ where $V = \{0, \ldots, m-1\}$ and $(i,j) \in E \Leftrightarrow \exists k : a_{i,k} \neq 0 \wedge a_{j,k} \neq 0$.

The undirected bipartite graph of A is defined as $G_b(A) = (\{r_0, \ldots, r_{m-1}\}, \{c_0, \ldots, c_{n-1}\}, E)$ with rows $R = \{r_i : i = 0, \ldots, m-1\}$ and columns $C = \{c_j : j = 0, \ldots, n-1\}$ and $(r_i, c_j) \in E \Leftrightarrow a_{i,j} \neq 0$.

The directed adjacency graph of an $n \times n$ matrix A is defined as $G_a(A) = (V, E)$ where $V = \{0, \ldots, n-1\}$ and $(i,j) \in E \Leftrightarrow a_{i,j} \neq 0$. The adjacency graph of symmetric matrices is undirected.

Outline

Consider an undirected graph $G = (V, E)$. Two vertices $i, j \in V$ are distance-k neighbors ($\delta(i,j) = k$) if the (edge-)length of a shortest path between i and j is less than or equal to k.

A (k,p)-coloring of G is a mapping $\kappa : V \rightarrow \{0, \ldots, p-1\}$ such that $\kappa(i) \neq \kappa(j)$ whenever $\delta(i,j) = k$.

The coloring is called restricted to W if only a subset of $W \subseteq V$ is involved.

Two columns in A are structurally orthogonal if and only if the corresponding vertices in $G_c(A)$ are not adjacent.

The (optimization version of the) COLUMN COMPRESSION problem is to find a $(1, p)$-coloring of $G_c(A)$ with minimal p.

Two rows in A are structurally orthogonal if and only if the corresponding vertices in $G_r(A)$ are not adjacent.

The (optimization version of the) ROW COMPRESSION problem is to find a $(1, p)$-coloring of $G_r(A)$ with minimal p.

Two columns in A are structurally orthogonal if and only if the corresponding vertices in $G_b(A)$ are non-distance-2 neighbors.

The (optimization version of the) COLUMN COMPRESSION problem is to find a $(2, p)$-coloring of $G_b(A) = (R, C, E)$ restricted to C with minimal p.

The (optimization version of the) ROW COMPRESSION problem is to find a $(2, p)$-coloring of $G_b(A) = (R, C, E)$ restricted to R with minimal p.

Sequential Distance-1 Coloring

E.g, Coloring G_c (G_r similar)

Distance-1 Coloring is NP-complete.

Heuristic: Sequential distance-1 coloring + ordering of vertices

$G_c(A) = (C, E)$
Forbidden $\in \boldsymbol{N}^p$
$\forall c \in C$
$\quad \forall c' : (c, c') \in E \wedge \textbf{color}(c') \neq 0$
$\qquad \textbf{Forbidden}[\textbf{color}(c')] = c$
$\quad \textbf{color}(c) = \min\{i : \textbf{Forbidden}[i] \neq c\}$

Forbidden keeps track of forbidden colors.

DISTANCE-2 COLORING is NP-complete.

Heuristic: Sequential distance-2 coloring + ordering of vertices

$G_b(A) = (R, C, E)$
Forbidden $\in \boldsymbol{N}^p$
$\forall c \in C$
$\quad \forall r : (c, r) \in E$
$\quad\quad \forall c' : (r, c') \in E \wedge \textbf{color}(c') \neq 0$
$\quad\quad\quad \textbf{Forbidden}[\textbf{color}(c')] = c$
$\quad \textbf{color}(c) = \min\{i : \textbf{Forbidden}[i] \neq c\}$

Consider an ordering of the vertices of a graph $G = (V, E)$ as

$$v_0, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_{n-1} \ .$$

- The **forward degree** of v_i is equal to the number of edges connecting v_i with $v_{i+1}, \ldots, v_{n-1}$.
- The **back(ward) degree** of v_i is equal to the number of edges connecting v_i with $v_0, \ldots, v_{i-1}$.
- The **degree** of v_i is equal to the sum of forward and back(ward) degrees.

Consider G_c for

$$\begin{pmatrix} x & & & & x & x \\ & x & x & x & & \\ & & x & & x & \\ x & & & x & & \\ & x & & & & x \end{pmatrix}$$

and vertex order

$$0, 1, 2, 3, 4, 5 \quad .$$

For vertex 2 we get

- forward degree $= 2$
- back(ward) degree $= 1$
- degree $= 3$

... in context of DISTANCE-1 COLORING.

Intuition: Order vertices imposing most restrictions first (new colors asap)

- Largest (Forward Degree) First
  For $i = 0, \ldots, n-1$: v_i has largest degree in $V \setminus \{v_0, \ldots, v_{i-1}\}$.
- Smallest (Forward Degree) Last
  For $i = n-1, \ldots, 0$: v_i has smallest degree in $V \setminus \{v_{n-1}, \ldots, v_{i+1}\}$.
- Incidence (Backward) Degree
  For $i = 0, \ldots, n-1$: v_i has largest degree in $\{v_0, \ldots, v_i\}$.

... to be combined (recursively) with tie-breaker (e.g. natural order)

Example: G_c

Largest First + Natural

For $i = 0, \dots, n-1$: v_i has largest degree in $V \setminus \{v_0, \dots, v_{i-1}\}$.

From

- all vertices have degree 3 in V
- 1 and 2 have degree 3 in $V \setminus \{0\}$
- 2 and 4 have degree 2 in $V \setminus \{0, 1\}$
- 4 and 5 have degree 1 in $V \setminus \{0, 1, 2\}$
- 3 and 5 have degree 0 in $V \setminus \{0, 1, 2, 4\}$
- 5 is sole vertex in $V \setminus \{0, 1, 2, 3, 4\}$

$$\begin{pmatrix} x & & & & x & x \\ & x & x & x & & \\ & & x & & x & \\ x & & & x & & \\ & x & & & & x \end{pmatrix}$$

follows

$$0, 1, 2, 4, 3, 5$$

Sequential coloring yields

$$0(a), 1(a), 2(b), 4(c), 3(c), 5(b)$$

Example: G_c

Incidence Degree + Natural

For $i = 0, \ldots, n-1$: v_i has largest degree in $\{v_0, \ldots, v_i\}$.

From

- all vertices i have degree 0 in $\{i\}$
- $i \in \{3,4,5\}$ have degree 1 in $\{0,i\}$
- $i \in \{1,2,4,5\}$ have degree 1 in $\{0,3,i\}$
- $i \in \{2,5\}$ have degree 2 in $\{0,3,1,i\}$
- $i \in \{4,5\}$ have degree 2 in $\{0,3,1,3,i\}$
- 5 is sole vertex in $V \setminus \{0,3,1,2,4\}$

$$\begin{pmatrix} x & & & & x & x \\ & x & x & x & & \\ & & x & & x & \\ x & & & x & & \\ & x & & & & x \end{pmatrix}$$

follows

$$0, 3, 1, 2, 4, 5$$

Sequential coloring yields

$$0(a), 3(b), 1(a), 2(c), 4(b), 5(c)$$

Example: G_c

Smallest Last + Reverse Natural

For $i = n-1, \ldots, 0$: v_i has smallest degree in $V \setminus \{v_{n-1}, \ldots, v_{i+1}\}$.

From

- all vertices have degree 3 in V
- 4, 1, and 0 have degree 2 in $V \setminus \{5\}$
- 0 has degree 1 in $V \setminus \{5, 4\}$
- 3, 2, and 1 have degree 2 in $V \setminus \{5, 4, 0\}$
- 2 and 1 have degree 1 in $V \setminus \{5, 4, 0, 3\}$
- 1 is sole vertex in $V \setminus \{5, 4, 0, 3, 2\}$

$$\begin{pmatrix} x & & & & x & x \\ & x & x & x & & \\ & & x & & x & \\ x & & & x & & \\ & x & & & & x \end{pmatrix}$$

follows

$$1, 2, 3, 0, 4, 5$$

Sequential coloring yields

$$1(a), 2(b), 3(c), 0(a), 4(c), 5(b)$$

Outline

Let $A = (a_{i,j}) = F'(\mathbf{x}) \in \boldsymbol{R}^{4\times 4}$.

$$A \cdot X^{(1)} = \begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} \\ a_{1,0} & a_{1,1} & & \\ a_{2,0} & & a_{2,2} & \\ a_{3,0} & & & a_{3,3} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{0,0} & \sum_{i=1}^{3} a_{0,i} \\ a_{1,0} & a_{1,1} \\ a_{2,0} & a_{2,2} \\ a_{3,0} & a_{3,3} \end{pmatrix}$$

$$Y_{(1)} \cdot A = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} \\ a_{1,0} & a_{1,1} & & \\ a_{2,0} & & a_{2,2} & \\ a_{3,0} & & & a_{3,3} \end{pmatrix} = \begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} \end{pmatrix}$$

All elements of A are pure in either $A \cdot X^{(1)}$ or $Y_{(1)} \cdot A$ yielding three derivative evaluations instead of n.

A star-coloring of the bipartite graph

1. uses disjoint color sets for rows and columns
2. allows for special neutral color of irrelevant vertices
3. is a distance-1 coloring
4. requires vertices that are adjacent to the same irrelevant vertex to have distinct colors
5. requires every path of (vertex-)length four to use at least three colors

Star-Coloring of Bipartite Graph

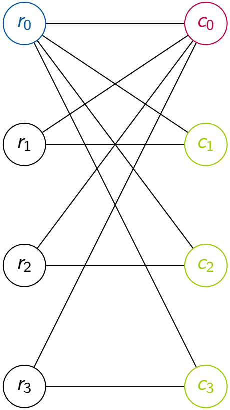
Example

$$\begin{pmatrix} a_{0,0} & a_{0,1} & a_{0,2} & a_{0,3} \\ a_{1,0} & a_{1,1} & & \\ a_{2,0} & & a_{2,2} & \\ a_{3,0} & & & a_{3,3} \end{pmatrix}$$

yields

$$X^{(1)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad Y^T_{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

or the other way around due to symmetry.



1. Disjoint color sets are used for the distinct column and row compression subproblems.
2. Certain rows/columns can be irrelevant as all their entries are pure in solutions to the corresponding column/row compression subproblem; e.g, r_1.
3. Edges represent nonzeros. End points of an edge must not carry the same color nor can both be irrelevant, e.g, (r_i, c_j) irrelevant $\Rightarrow a_{i,j}$ potentially not pure.
4. If two vertices adjacent to the same irrelevant vertex had the same color then the corresponding matrix entries cannot be pure, e.g, c_0 and c_2 same color and adjacent to $r_2 \Rightarrow a_{2,0} + a_{2,2}$.
5. ...

Star-Coloring of Bipartite Graph

Proof: Correctness of 5.

W.l.o.g. consider $\binom{4}{2} = 6$ two-colorings of path (r_i, c_j, r_k, c_l), namely

1. o o o o
2. o o o o
3. o o o o
4. o o o o
5. o o o o
6. o o o o

1.-4. are distance-1 colorings and 5. and 6. are symmetric, yielding inability to recover entry (k, j) from

$$(1, \ldots, 1) \cdot \begin{pmatrix} a_{i,j} & & \\ \vdots & & \\ a_{k,j} & \ldots & a_{k,l} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a_{i,j} & & \\ \vdots & & \\ a_{k,j} & \ldots & a_{k,l} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Proof: Sufficiency of 5.

Consider $\binom{4}{2} \cdot 2! = 12$ three-colorings of path (r_i, c_j, r_k, c_l).

- 6 are distance-1 colorings
- 6 are pair-wise symmetric
- investigation of remaining 3 scenarios, namely
  1. o o o o ($\Rightarrow a_{k,j}$ pure in c_j)
  2. o o o o ($\Rightarrow a_{k,j}$ pure in r_k)
  3. o o o o ($\Rightarrow$ same color for row and column $\Rightarrow$ neutral; $a_{k,j}$ pure in both c_j and r_k)

  yields
  1. o not neutral
  2. o not neutral
  3. o neutral

$$\begin{pmatrix} a_{i,j} & & \\ \vdots & & \\ a_{k,j} & \dots & a_{k,l} \end{pmatrix}$$

and ability to recover all relevant Jacobian entries. q.e.d.

See literature for heuristics for star-coloring.

Outline

ColPack

https://github.com/CSCsw/ColPack

implements a range of coloring methods for Jacobian compression.

See A. Gebremedhin et al.: What Color is your Jacobian? SIAM, 2005 further details.

Outline

Summary

- Introduction to direct Jacobian compression based on graph coloring
- Formulation of Jacobian compression problems as graph coloring problems
- Heuristics for distance-1 graph coloring

Next Steps

- Practice derivation of graph representations of matrices.
- Practice application of heuristics for distance-1 graph coloring.
- Continue the course to find out more ...