Jacobians, Graphs, Combinatorics

Uwe Naumann

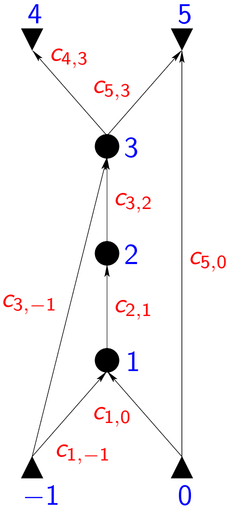
LuFG Software and Tools for Computational Engineering
Dept. of Computer Science, RWTH Aachen University, Germany

Content

(Linearized) Computational Graph

$t = x_0 \cdot \sin(x_0 \cdot x_1)$
$x_0 = \cos(t)$
$x_1 = t/x_1$

$v_{-1} = x_0$
$v_0 = x_1$
$v_1 = v_{-1} \cdot v_0 \qquad c_{1,-1} = v_0;\ c_{1,0} = \dots$
$v_2 = \sin(v_1) \qquad c_{2,1} = \cos(v_1)$
$v_3 = v_{-1} \cdot v_2 \qquad c_{3,-1} = v_2;\ c_{3,2} = \dots$
$v_4 = \cos(v_3) \qquad c_{4,3} = -\sin(v_3)$
$v_5 = v_3/v_0 \qquad c_{5,3} = 1/v_0;\ c_{5,0} = \dots$
$x_0 = v_4$
$x_1 = v_5$



$G = (V, E)$

Elimination of Paths

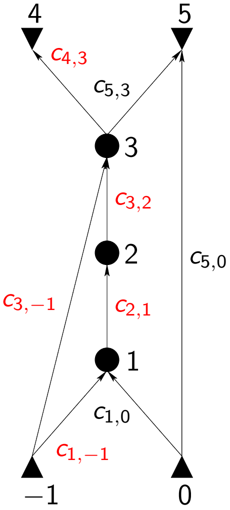
$$\frac{\partial v_k}{\partial v_l} \equiv f'_{k,l} = \sum_{[l \rightarrow k]} \prod_{(i,j) \in [l \rightarrow k]} c_{j,i}$$

For example,

$$f'(4,-1) = c_{3,-1} \cdot c_{4,3} + c_{1,-1} \cdot c_{2,1} \cdot c_{3,2} \cdot c_{4,3}$$

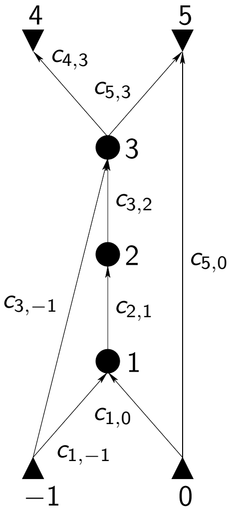
$$= (c_{3,-1} + c_{1,-1} \cdot c_{2,1} \cdot c_{3,2}) \cdot c_{4,3}$$

Objective ...



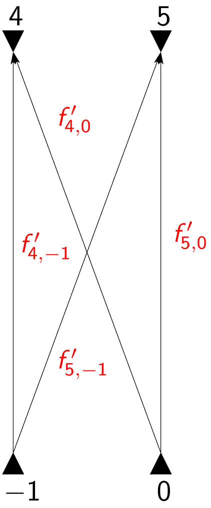
- W. Baur and V. Strassen: *The complexity of partial derivatives.* 1983

Bipartite Computational Graph



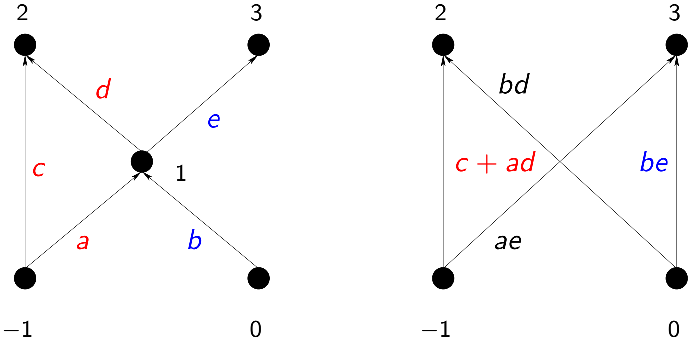
G

$\Longrightarrow$



G'

Elimination of Vertices



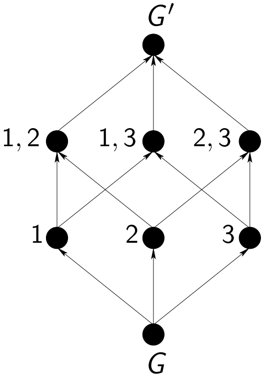
$$\text{Cost}(j) = |P_j| \cdot |S_j| \quad (\text{Cost}(1) = |P_1| \cdot |S_1| = 2 \cdot 2 = 4)$$

- A. Griewank and S. Reese: *On the calculation of Jacobian Matrices by the Markovitz rule.* Proceedings of AD1991, SIAM (1991)
- K. Herley: *On minimal fill-in Jacobian accumulation.* ANL, 1992

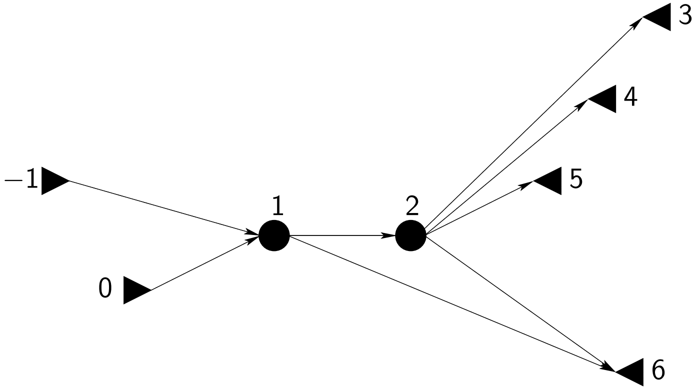
Search Space (M_v)

For example, $|V| = 3$:

- $\longrightarrow$ $|V|!$ elimination sequences
- $\longrightarrow$ shortest path problem in cost-enhanced *metagraph* M_v
- $\longrightarrow$ size of M_v is exponential in size of G



Lion Graph



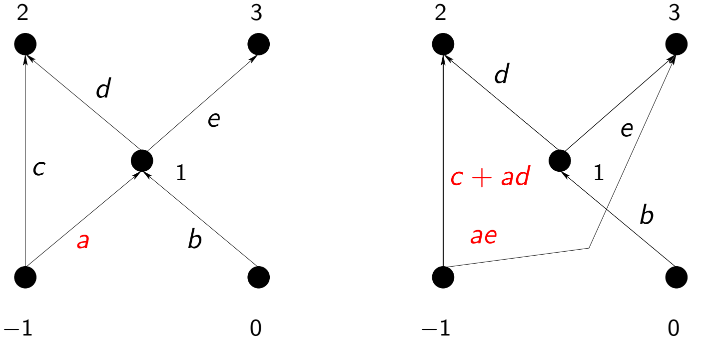
$$f'_{3,-1} = c_{1,-1} \cdot c_{2,1} \cdot c_{3,2}$$

$$\ldots$$

$$f'_{6,-1} = c_{1,-1} \cdot c_{2,1} \cdot c_{6,2} + c_{1,-1} \cdot c_{6,1} = c_{1,-1} \cdot (c_{2,1} \cdot c_{6,2} + c_{6,1})$$

$$f'_{6,0} = c_{1,0} \cdot c_{2,1} \cdot c_{6,2} + c_{1,0} \cdot c_{6,1} = c_{1,0} \cdot (c_{2,1} \cdot c_{6,2} + c_{6,1})$$

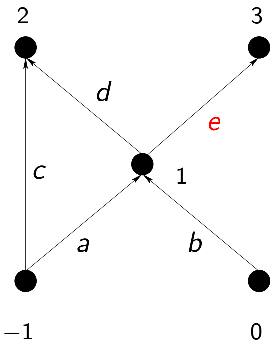
(Front-)Elimination of Edges

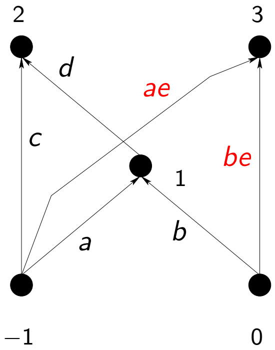


$$\text{Cost}((i,j)) = |S_j| \quad (\text{Cost}((-1,1)) = |S_1| = 2)$$

- U. Naumann: *Elimination Techniques for Cheap Jacobians.* Proceedings of AD2000, Springer (2000)

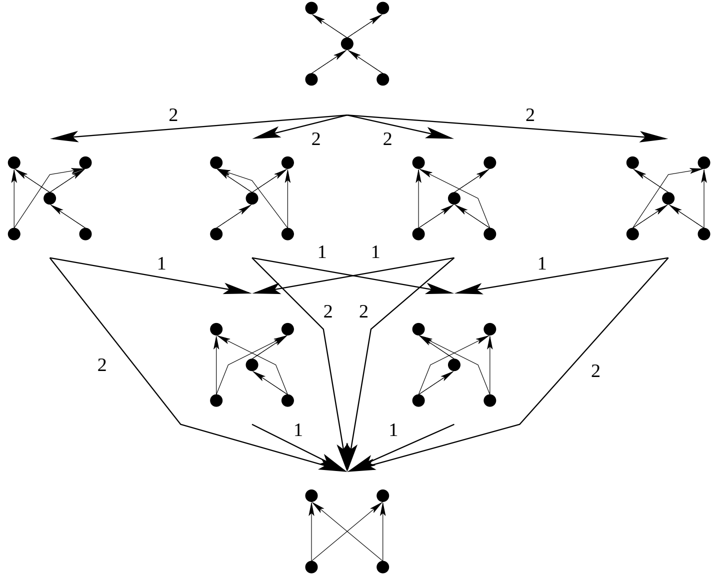
(Back-)Elimination of Edges



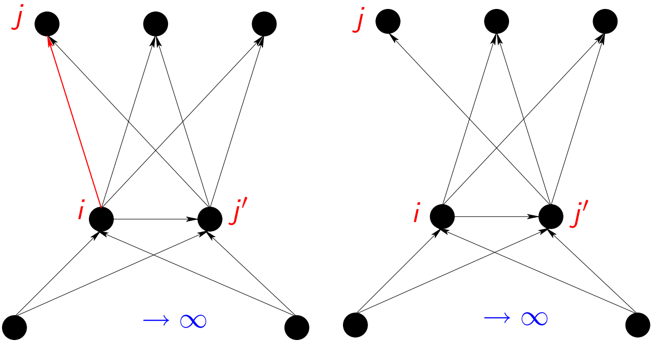


$$\mathrm{Cost}((i,j)) = |P_i| \quad (\mathrm{Cost}((1,3)) = |P_1| = 2)$$

Search Space (M_e)

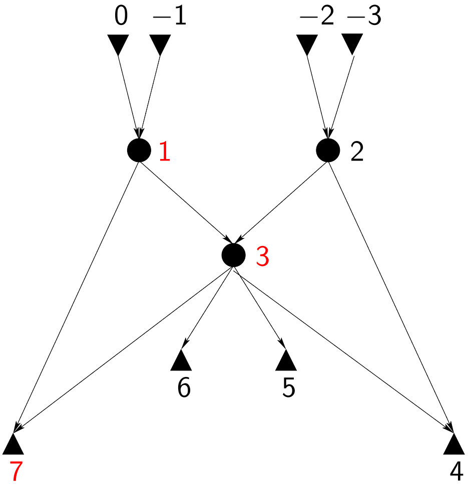


Rerouting

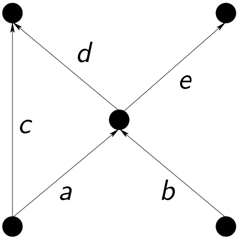


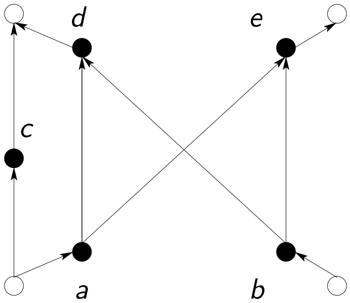
- A. Griewank and O. Vogel: *Analysis and exploitation of Jacobian scarcity.* Proceedings of HPSC (2003)

Bat Graph



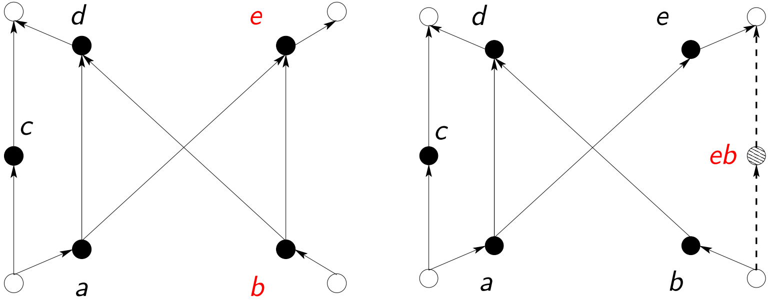
Dual Computational Graph





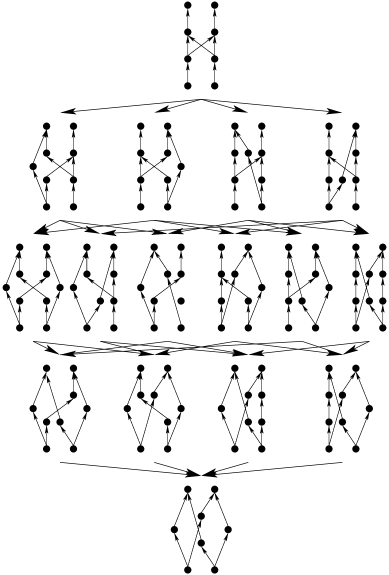
- U. Naumann: *Optimal Accumulation of Jacobian matrices by elimination methods on the dual computational graph.* Math. Prog., Springer (2004)

Elimination of Faces



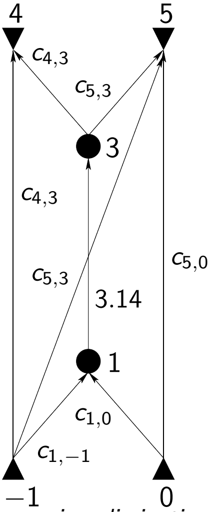
$$\text{Cost}((i,j)) = 1$$

Search Space (M_f)



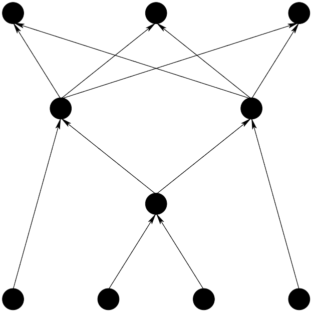
Compile-time Elimination

$$t = x_0 + (x_0 \cdot x_1) \cdot 3.14$$
$$x_0 = \cos(t)$$
$$x_1 = t/x_1$$



- U. Naumann and J. Utke: *Optimality-preserving elimination of linearities in Jacobian accumulation.* Electronic Transactions on Numerical Analysis, KSU (2005).

Heuristics



- U. Naumann: *An Enhanced Markovitz Rule for Accumulating Jacobians Efficiently*. Proceedings of 15th Conference on Scientific Computing (ALGORITHMY 2000).

Dynamic Programming

$$F'(x) = Q_m \left[\prod_{j=1}^{l+m} \prod_{i \prec j} C_{ji} \right] P_n^T \in \mathbb{R}^{m \times n}$$

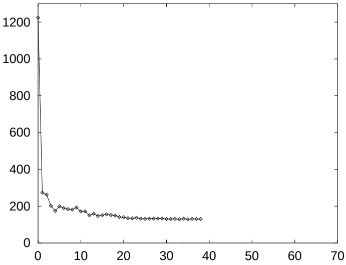
where

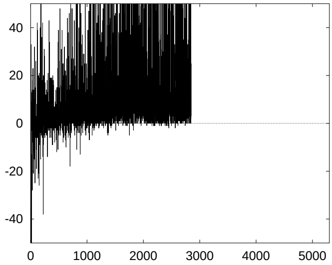
$$P_n \equiv \left[I_n, 0 \right] \in \mathbb{R}^{n \times q} \quad \text{and} \quad Q_m \equiv \left[0,\ I_m \right] \in \mathbb{R}^{m \times q}$$

- A. Griewank and U. Naumann: *Accumulating Jacobians as Chained Sparse Matrix Products*. Math. Prog., Springer (2003).

Simulated Annealing

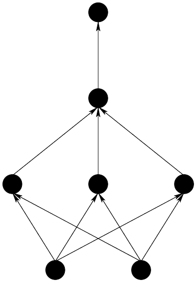
- start sequence, start temperature
- (logarithmic) cooling schedule
- acceptance of worse sequence with Metropolis probability
- run for as long as you like ...





▶ U. Naumann: *Cheaper Jacobians by Simulated Annealing.* SIAM J. Opt., SIAM (2002).

Single-Expression-Use Graphs



- U. Naumann and Y. Hu: *Optimal Vertex Elimination in Single-Expression-Use Graphs.* AIB-2006-08, RWTH Aachen (2006).

Proving NP-completeness

1. pick known NP-complete problem (NPP)
2. derive polynomially instance of your problem for each instance of NPP
3. verify given solution in polynomial time

▶ U. Naumann: Optimal Jacobian Accumulation is NP-complete. Under review by Math. Prog., Springer

Ensemble Computation

Given a collection $C = \{C_\nu \subseteq A : \nu = 1, \ldots, |C|\}$ (Jacobian) of subsets $C_\nu = \{c_i^\nu : i = 1, \ldots, |C_\nu|\}$ (Jacobian entries) of a finite set A (elemental partial derivatives) and a positive integer Ω (max. nr. of scalar multiplications) is there a sequence $u_i = s_i \cup t_i$ (scalar multiplications) for $i = 1, \ldots, \omega$ of $\omega \leq \Omega$ union operations, where each s_i and t_i is either $\{a\}$ (elemental partial derivative) for some $a \in A$ or u_j (previously accumulated partial derivative) for some $j < i$, such that s_i and t_i are disjoint for $i = 1, \ldots, \omega$ and such that for every subset $C_\nu \in C$, $\nu = 1, \ldots, |C|$, there is some u_i, $1 \leq i \leq \omega$, that is identical to C_ν (all Jacobian entries are computed).

Theorem

EC is NP-complete.

M. Garey and D. Johnson. Computers and Intractability. 1979

EC Example

Let an instance of EC be given by

$$A = \{a_1, a_2, a_3, a_4\}$$
$$C = \{\{a_1, a_2\}, \{a_2, a_3, a_4\}, \{a_1, a_3, a_4\}\}$$

and $\Omega = 4$. The answer to the decision problem is positive with a corresponding instance given by

$$\begin{aligned} C_1 = u_1 &= \{a_1\} \cup \{a_2\} \\ u_2 &= \{a_3\} \cup \{a_4\} \\ C_2 = u_3 &= \{a_2\} \cup u_2 \\ C_3 = u_4 &= \{a_1\} \cup u_2 \quad . \end{aligned}$$

Optimal Jacobian Accumulation

Given a linearized computational graph G of a vector function F and a positive integer Ω is there a sequence of scalar assignments $u_k = s_k \circ t_k,\ \circ \in \{+, *\},\ k = 1, \ldots, \omega,$ where each s_k and t_k is either $c_{j,i}$ for some $(i,j) \in E$ or $u_{k'}$ for some $k' < k$ such that $\omega \leq \Omega$ and for every Jacobian entry there is some identical $u_k,\ k \leq \omega$?

Example: *Lion*

$c_{6,1} := c_{6,1} + c_{6,2}c_{2,1};\ c_{2,-1} = c_{2,1}c_{1,-1};\ c_{2,0} = c_{2,1}c_{1,0}$
$c_{6,-1} = c_{6,1}c_{1,-1};\ c_{6,0} = c_{6,1}c_{1,0};\ c_{3,-1} = c_{3,2}c_{2,-1}$
$c_{3,0} = c_{3,2}c_{2,0};\ c_{4,-1} = c_{4,2}c_{2,-1};\ c_{4,0} = c_{4,2}c_{2,0}$
$c_{5,-1} = c_{5,2}c_{2,-1};\ c_{5,0} = c_{5,2}c_{2,0}$

Reduction EC → OJA

Consider $\mathbf{y} = F(\mathbf{x}, \mathbf{a})$ where $\mathbf{x} \in \mathbb{R}^{|C|}$, $\mathbf{a} \in \mathbb{R}^{|A|}$ is a vector containing all elements of A, and $F : \mathbb{R}^{|C|+|A|} \rightarrow \mathbb{R}^{|C|}$ defined as

$$y_\nu = x_\nu * \prod_{j=1}^{|C_\nu|} c_j^\nu$$

for $\nu = 1, \ldots, |C|$ and where c_j^ν is equal to some $a \in A$ for all ν and j. This transformation is linear with respect to the original instance of Ensemble Computation in both space and time. The Jacobian $F'(\mathbf{x}, \mathbf{a})$ is a diagonal matrix with nonzero entries

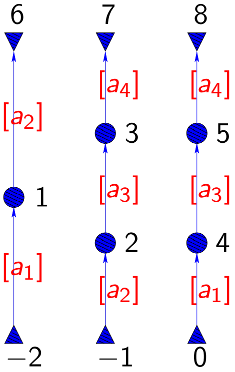
$$f_{\nu,\nu} = \prod_{j=1}^{|C_\nu|} c_j^\nu$$

for $\nu = 1, \ldots, |C|$.

Example

$A = \{a_1, a_2, a_3, a_4\}$
$C = \{\{a_1, a_2\}, \{a_2, a_3, a_4\}, \{a_1, a_3, a_4\}\}$

$f'_{1,1} = c_{6,-2} = a_1 * a_2$
$u_2 = c_{7,2} = c_{8,4} = a_3 * a_4$
$f'_{2,2} = c_{7,-1} = a_2 * u_2$
$f'_{3,3} = c_{8,0} = a_1 * u_2$



EC $\Leftrightarrow$ OJA

$\Leftarrow$: Simply subsitute $*$ for $\cup$.

$\Rightarrow$:
1. No additions; simply subsitute $*$ for $\cup$.
2. Suppose that there is some $i \leq \omega$ such that $s_i \cap t_i = \{b\}$. Hence the computation of u_i in the Jacobian accumulation code involves a factor $b * b$. Note that such a factor is not part of any Jacobian entry which implies that the computation of u_i is obsolete and therefore cannot be part of an optimal Jacobian accumulation code.

Consequences

1. "Rows and columns" of F' are NP-complete.

$$y = \sum_{\nu=1}^{|C|} y_\nu = \sum_{\nu=1}^{|C|} \left(x_\nu * \prod_{j=1}^{|C_\nu|} c_j^\nu \right) \quad .$$

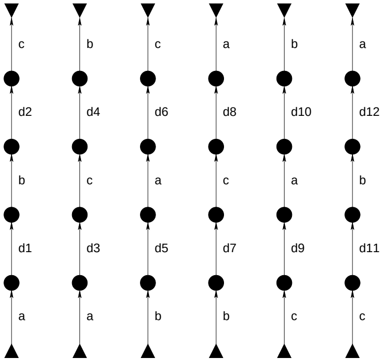
2. "Tangents and adjoints" are NP-complete.

$$y_\nu = x * \dot{x}_\nu * \prod_{j=2}^{|C_\nu|} c_j^\nu \quad .$$

3. "Partial derivatives of arbitrary order" are NP-complete.

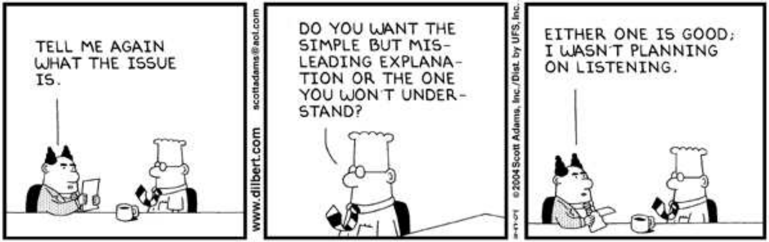
$$y_\nu = \frac{x_\nu^q}{q!} \prod_{j=1}^{|C_\nu|} c_j^\nu \quad .$$

Discussion



$$OPS(F) = 6 * 5 = 30$$
$$OPS(F') = 2 + 6 * 2 = 14$$

Conclusion


TELL ME AGAIN WHAT THE ISSUE IS.
scottadams@aol.com
www.dilbert.com
DO YOU WANT THE SIMPLE BUT MIS-LEADING EXPLANA-TION OR THE ONE YOU WON'T UNDER-STAND?
© 2004 Scott Adams, Inc./Dist. by UFS, Inc.
EITHER ONE IS GOOD; I WASN'T PLANNING ON LISTENING.