Optimal Jacobian Accumulation is NP-complete

Uwe Naumann[1]

Abstract

We show that the problem of accumulating Jacobian matrices by using a minimal number of floating-point operations is NP-complete by reduction from Ensemble Computation. The proof makes use of the fact that, deviating from the state-of-the-art assumption, algebraic dependences can exist between the local partial derivatives. It follows immediately that the same problem for directional derivatives, adjoints, and higher derivatives is NP-complete, too.

We consider the automatic differentiation (AD) [2] of an implementation of a non-linear vector function $\mathbf{y} = F(\mathbf{x}, \mathbf{a}), F : I\!R^{n+\tilde{n}} \supseteq D \rightarrow I\!R^m$, as a computer program.[2] With the Jacobian matrix F' of F defined as usual *tangent-linear* ($\dot{\mathbf{y}} = F'(\mathbf{x}, \mathbf{a}) * \dot{\mathbf{x}}$, $\dot{\mathbf{x}} \in I\!R^n$) and *adjoint* ($\bar{\mathbf{x}} = (F'(\mathbf{x}, \mathbf{a}))^T * \bar{\mathbf{y}}$, $\bar{\mathbf{y}} \in I\!R^m$) versions of numerical simulation programs with potentially complicated intra- and inter-procedural flow of control can be generated automatically by AD tools. This technique has been proved extremely useful in the context of numerous applications of computational science and engineering requiring numerical methods that are based on derivative information. For the purpose of this paper we may assume trivial flow of control in the form of a straight-line program. Similarly, one may consider the evaluation of an arbitrary function at a given point to fix the flow of control.

Our interest lies in the computation of the Jacobian of the *active* outputs (or *dependent* variables) $\mathbf{y} = (y_j)_{j=1,\ldots,m}$ with respect to the *active* inputs (or *independent* variables) $\mathbf{x} = (x_i)_{i=1,\ldots,n}$. The $\tilde{n}-$vector $\mathbf{a}$ contains all *passive* inputs. Conceptually, AD decomposes the program into a sequence of scalar assignments $v_j = \varphi_j(v_i)_{i \prec j}$ for $j = 1, \ldots, p+m$, where we follow the notation in [2]. We refer to this equation as the *code list* of F, and we set $x_i = v_{i-n}$ for $i = 1, \ldots, n$ and $y_j = v_{p+j}$ for $j = 1, \ldots, m$. The v_j, $j = 1, \ldots, p$, are referred to as *intermediate* variables. The notation $i \prec j$ marks a direct dependence of v_j on v_i meaning that v_i is an argument of the *elemental function*[3] φ_j. The code list induces a directed acyclic graph $G = (V, E)$ such that $V = \{1-n, \ldots, p+m\}$ and $(i, j) \in E \Leftrightarrow i \prec j$. Assuming that all elemental functions are continuously differentiable at their respective arguments all local partial derivatives can be computed by a single evaluation of the *linearized* code list

$$c_{j,i} = \frac{\partial \varphi_j}{\partial v_i}(v_k)_{k \prec j} \quad \forall i \prec j$$
$$v_j = \varphi_j(v_i)_{i \prec j}$$

[1]Software and Tools for Computational Engineering, Department of Computer Science, RWTH Aachen University, 52056 Aachen, Germany, `http://www.stce.rwth-aachen.de`, `naumann@stce.rwth-aachen.de`

[2]F is used to refer to the given implementation.

[3]Elemental functions are the arithmetic operators and intrinsic functions provided by the programming language.

for $j = 1, \ldots, p+m$ and for given values of $\mathbf{x}$ and $\mathbf{a}$. The corresponding linearized version of G is obtained by attaching the $c_{j,i}$ to the corresponding edges (i,j). Various elimination techniques have been proposed for efficient Jacobian accumulation based on G [3].

Theorem 1 Optimal Jacobian Accumulation *is NP-complete.*

We reduce from Ensemble Computation [1]. A given solution is verified in polynomial time by counting the number of operations.

Given an arbitrary instance of Ensemble Computation we define the corresponding Optimal Jacobian Accumulation problem as follows:

Consider $\mathbf{y} = F(\mathbf{x}, \mathbf{a})$ where $\mathbf{x} \in I\!R^{|C|}$, $\mathbf{a} \equiv (a_j)_{j=1,\ldots,|A|} \in I\!R^{|A|}$ is a vector containing all elements of A, and $F : I\!R^{|C|+|A|} \rightarrow I\!R^{|C|}$ 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_i, $i = 1, \ldots, |A|$, for all ν and j. The elements of A are set to be random numbers. 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 $\nu = 1, \ldots, |C|$. Is there a Jacobian accumulation code for $F'(\mathbf{x}, \mathbf{a})$ of length less than Ω? We claim that the answer is positive if and only if there is a solution of the corresponding Ensemble Computation problem.

The presentation is based on [4]. An outline of the proof will be presented. The result can be used to show that the optimal accumulation of directional derivatives as well as of adjoints and of higher derivatives are NP-complete too.

References

[1] M. Garey and D. Johnson. *Computers and Intractability - A Guide to the Theory of NP-completeness.* W. H. Freeman and Company, 1979.

[2] A. Griewank. *Evaluating Derivatives. Principles and Techniques of Algorithmic Differentiation.* Number 19 in Frontiers in Applied Mathematics. SIAM, Philadelphia, 2000.

[3] U. Naumann. Optimal accumulation of Jacobian matrices by elimination methods on the dual computational graph. *Math. Prog.*, 3(99):399–421, 2004.

[4] U. Naumann. Optimal Jacobian accumulation is NP-complete. *Math. Prog.*, 2006. To appear.