Optimal Bi-directional Determination of Sparse Jacobian Matrices[1]

Mini Goyal and **Shahadat Hossain**
Department of Mathematics and Computer Science,
University of Lethbridge, Canada

Given "seed" matrix $S \in R^{n \times p}$, the Jacobian matrix $J \in R^{m \times n}$ of a mapping $F : R^n \to R^m$ at $x \in R^n$ can be "compressed" by computing the product JS using Automatic Differentiation (AD) techniques. The nonzero entries of matrix J can be recovered (i.e. restored) by solving for them in the linear system of equations

$$JS = B$$

where B is the compressed Jacobian matrix obtained e.g., via AD forward mode. Direct methods allow the nonzero entries to be "read off" the compressed Jacobian. A lower bound on the number of matrix-vector products to completely determine J directly is given by the maximum number of nonzero entries in any row of J. For matrices with few dense rows and columns the one-directional compression (i.e column compression) may not be able to exploit the known sparsity satisfactorily. Hossain and Steihaug [4], and Coleman and Verma [1] independently proposed heuristic techniques where the matrix is compressed both in column direction and in row direction (via reverse AD). More recently Gebremedhin, Manne, and Pothen [3] have proposed a greedy approach termed *distance 3/2 bi-coloring scheme* that exploits known sparsity in both column and row directions to determine Jacobian matrices. While in one-directional compression a satisfactory lower bound on the number of matrix-vector products (and vector-matrix products) is obtained easily, similar lower bounds for bi-directional compression are not known. In this paper we present optimal bi-directional determination of sparse Jacobian matrices using integer linear programming (ILP) methods. We propose an ILP model of the underlying combinatorial problem and discuss implementation issues. Numerical tests are being conducted using ILOG CPLEX MIP optimizer on selected problems from Harwell-Boeing[2] test collection. Preliminary results are promising.

An ILP formulation of the bi-directional determination of sparse Jacobian matrices

Let $A \in R^{m \times n}$ be a matrix with known sparsity pattern. Define bipartite graph

$$G_b(A) = (U \cup V, E)$$

where U and V are the sets of vertices corresponding to the columns and rows of A respectively, and E is the set of edges between vertices $u_i \in U$ and $v_j \in V$ such that $\{u_i, v_j\} \in E$ if $a_{ij} \neq 0$. The p-coloring of graph $G = (V, E)$ is a function $\Phi : V \to \{1, \ldots, p\}$ such that $\phi(v_i) \neq \phi(v_j)$ if $\{v_i, v_j\} \in E$. A mapping ϕ is called *bi-directional p-coloring* [4] if the following conditions apply:

1. ϕ is a p-coloring,
2. the set of colors used on vertices in U and V are disjoint, i.e., for $u_i \in U$ and $v_j \in V$ $\phi(u_i) \neq \phi(v_j)$, and
3. every path of length 3 in $G_b(A)$ uses at least three different colors.

The bi-chromatic number of $G_b(A)$ is the smallest p for which $G_b(A)$ has a bi-directional p-coloring. For brevity, henceforth, we will omit p and use "bi-coloring" instead of *bi-directional p-coloring* when there are no confusion. Finding a bi-coloring with minimum number of colors has been shown to be

[1]This research is supported in part by the Natural Sciences and Engineering Research Council of Canada.

equivalent to determining seed matrices V and W such that the number of matrix-vector products AV and the vector-matrix products $W^T A$ is minimized in determining matrix A directly [4].

The ILP presented in Figure 1 models the optimal bi-coloring of $G_b(A)$ where w_j indicates

$$
\begin{array}{ll}
\text{minimize } \sum_{j=1}^{p_U+p_V} w_j & \\
\text{subject to} & \\
\sum_{j=1}^{p_U+p_V} x_{i,j} = 1 & \text{for } i \in U \cup V \\
x_{p,j} + x_{q,j'} + x_{r,j} + x_{s,j'} \leq (w_j + w_{j'} + 1) & \text{for each 3-path } p-q-r-s \\
 & \text{and for each color pair } \{j, j'\} \\
 & 1 \leq j \leq p_U, \ \ p_U + 1 \leq j' \leq p_U + p_V \\
w_j \leq \sum_{i \in U \cup V} x_{i,j} & \text{for } j = 1, \ldots, p_U + p_V \\
\sum_{i \in U} x_{i,j} \leq n w_j & \text{for } j = 1, \ldots, p_U \\
\sum_{i \in V} x_{i,j} \leq m w_j & \text{for } j = p_U + 1, \ldots, p_U + p_V
\end{array}
$$

Figure 1: Bi-ILP – an ILP model for computing optimal bi-coloring.

whether (= 1) or not (= 0) color j has been assumed by some vertex, x_{ij} indicates whether (= 1) or not (= 0) vertex i has assumed color j. The quantities p_U and p_V denote upper bounds on the number of colors we allow for the column and the row vertices respectively.

References

[1] T. F. Coleman and A. Verma. The efficient computation of sparse Jacobian matrices using automatic differentiation. *SIAM J. Sci. Comput.*, 19(4):1210–1233, 1998.

[2] I. S. Duff, R. G. Grimes, and J. G. Lewis. Sparse matrix test problems. *ACM Transactions on Mathematical Software*, 15(1):1–14, 1989.

[3] A. H. Gebremedhin, F. Manne, and A. Pothen. What color is your Jacobian? graph coloring for computing derivatives. Technical Report, 2004. Accepted by SIAM Review.

[4] A. S. Hossain and T. Steihaug. Computing a sparse Jacobian matrix by rows and columns. *Optimization Methods and Software*, 10:33–48, 1998.