ショートノート

高速微分法における変数消去のグラフ論的考察†

室田 一雄††　久保田 光一††

高速微分法の計算を効率化するために計算グラフを適当に縮約したりブロックに分けたりして中間変数を消去するという自然な着想がある．特に，多変数のベクトル値関数のヤコビ行列を計算するには，多変数のスカラ値関数に対する算法を関数ごとに繰り返し独立に適用するという以上の根本的工夫は知られていないので，この計算の手間を少しでも減らすために各関数に共通な計算を検出することが望ましい．本稿ではこのような技法に関連したグラフ論的な事実を整理し，計算グラフの部分グラフでスカラ値の関数副プログラムと見なせる構造をもつものを抽出するための効率的算法を与える．

1. まえがき

高速微分法（文献 4)およびその文献表参照）は今やその重要性が広く認識され，実用化に向けての研究が行われている．本稿は，計算グラフの縮約，ブロック化などに基づいて高速微分法の実際的計算効率を高めるためのグラフ論的技法に関して基本的事実を整理することを主な目的とする．用語は文献 4)に従う．

計算過程が条件分岐などを含まず計算グラフの構造が入力データによらずに定まる“静的な”場合には，あらかじめ計算グラフの構造を調べて前処理を施しておくことにより，実行時の計算効率を高めることができる．特に，加減算に対応する要素的偏導関数の値が入力データによらないことに着目して計算グラフを簡約化する試みが文献 8)に報告されている．

計算グラフの構造が入力データに応じて定まる“動的な”場合に計算グラフの簡約化のために多くの手間を費やすことは実際的でない．しかし，n 変数 $x=(x_1,\cdots,x_n)$ のベクトル値関数 $f=(f_1,\cdots,f_m)$ のヤコビ行列 $(\partial f_i/\partial x_j)$ を求める際には，同一の計算グラフに対して高速微分法の算法を繰り返し適用する（詳しくは文献 4)参照）ので計算グラフの簡約化が有効である場合もある．

例えば，多くの中間変数を含む複雑な計算過程で表される 1 変数スカラ値関数 g を含む $[v_1:=x_1\cdot x_2;\ v_2:=g(v_1);\ f_1:=x_1\cdot v_2;\ f_2:=x_2\cdot v_2]$ という計算過程（その計算グラフ G を図 1（a）に示す）で定義される $f_1(x_1,x_2)$, $f_2(x_1,x_2)$ のヤコビ行列を計算する際には，$(\partial f_1/\partial x_1, \partial f_1/\partial x_2)$ と $(\partial f_2/\partial x_1, \partial f_2/\partial x_2)$ とを別々に計算するよりも，まず g を表す部分グラフに対して高速微分法の算法を適用して $\partial g/\partial v_1$ を求め，G から g の部分に含まれる中間変数を消去して v_1 から v_2 への枝を付け加えたグラフ G'（図 1 (b)）を作った後に通常の算法を適用する方が有利である．このような変数消去による計算グラフの簡略化は線形計算など計算グラフの構造が既知の場合には特に有効である[7]．

二つの計算グラフ（計算過程）が同じ初等関数を表現するかどうかを判定することは一般に不可能であること[6]から考えて，一つの関数を表現する計算グラフ（複数ありうる）の標準形を適当に定義してそれを求めることはできない．さらに，高速微分法は中間変数を陽に扱うことによって計算過程のもつ疎性を自然な形で利用することに成功している算法であるから，変数を消去して計算グラフを簡約化して効率を高めようとする方向性は，基本的に高速微分法の精神と相いれ

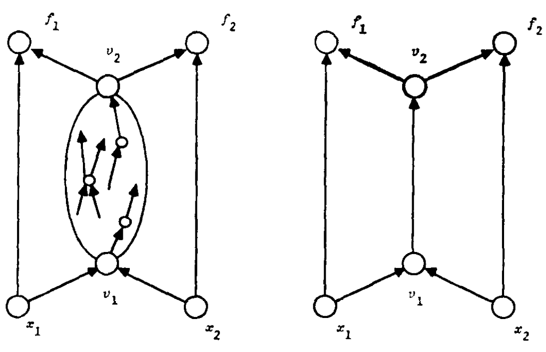

図 1　計算グラフの縮約
Fig. 1　Elimination of vertices in a computational graph.

† On Elimination of Intermediate Variables in Fast Automatic Differentiation by KAZUO MUROTA and KOICHI KUBOTA (Department of Mathematical Engineering and Information Physics, Faculty of Engineering, University of Tokyo).

†† 東京大学工学部計数工学科

ない．したがって，本稿は考察するような手法は高速微分法にとってあくまで補助的な実用上の工夫と考えるべきであろう．

2. 計算グラフと縮約可能な部分グラフ

n 入力変数 $(x_1,\cdots,x_n)$ のベクトル値関数 $f=(f_1,\cdots,f_m)$ の計算グラフは，入力変数，中間変数（f_i も含める），定数に対応する頂点集合 V と計算ステップの関数関係に対応する枝集合 E をもつ無閉路（非巡回）グラフ $G=(V,E)$ である．枝 $e\in E$ の始点を ∂^+e $(\in V)$，終点を $\partial^-e(\in V)$ と表す．次の記号を用いる．

$$\delta^+v=\{e\in E\,|\,\partial^+e=v\},\quad v\in V,$$
$$\delta^-v=\{e\in E\,|\,\partial^-e=v\},\quad v\in V,$$
$$X(G)=\{v\in V\,|\,\delta^-v=\phi\},$$
$$F(G)=\{v\in V\,|\,\delta^+v=\phi\}.$$

$u\in V$ から $v\in V$ に至る有向道が存在することを $u\rightarrow v$ と書く．特に $v\rightarrow v$ である．

頂点 v は $X(G)$，$F(G)$ に属するときそれぞれ極小，極大であるといわれる．本稿では

$$X(G)=\{x_1,\cdots,x_n\},\quad F(G)=\{f_1,\cdots,f_m\}$$

と仮定する．これは頂点 $v\in V$ のうちである x_j とある f_i に対して $x_j\rightarrow v\rightarrow f_i$ であるようなものだけを取り出して考えることと同等である．

周知のように

$$\partial f_i/\partial x_j=\sum d(p) \qquad (1)$$

である．ただし，和は x_j から f_i に至る G 上の有向道 p に関してとり，$d(p)$ は p 上の枝に付随する要素的偏導関数値の積を表す．

$G=(V,E)$ の部分グラフ $G'=(V',E')$ で，

i）$\forall v\in V$: $\delta^+v\cap E'=\phi$ あるいは $\delta^+v\subset E'$,

ii）$\forall v\in V$: $\delta^-v\cap E'=\phi$ あるいは $\delta^-v\subset E'$

を満たすものを**可縮部分グラフ**と定義する（δ^+，δ^- は G における接続関係を表す）．特に，$|F(G')|=1$ である可縮部分グラフ G' を**傘**，$|F(G')|=|X(G')|=1$ である可縮部分グラフ G' を**閉傘**と呼ぶ．

G' が可縮部分グラフのとき，$u\in X(G')$ から $v\in F(G')$ に至る G 上の有向道はすべて G' に含まれるので，このような有向道に沿う要素的偏導関数の積をあらかじめ計算しておけば (1) 式を $i=1,\cdots,m$; $j=1,\cdots,n$ に対して計算する際の手間が節約できる．このような前処理は，G において $V'-(X(G')\cup F(G'))$ の頂点およびそれに接続する枝を除去し，$u\rightarrow v$ であるすべての $u\in X(G')$，$v\in F(G')$ に対して枝 (u,v) を付加することに相当する．これを可縮部分グラフ G' の**縮約**と呼ぶ．G' の縮約の効果は

$\varDelta(G')$＝除去された枝の数－付加された枝の数

で見積れる．一般には $\varDelta(G')$ の符号は定まらないが，G' が傘の場合には $\varDelta(G')\geqq 0$ である．

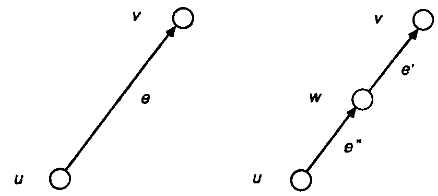

図 2 変数の複写による頂点の追加

Fig. 2 New vertex created by copying a variable.

計算グラフ $G=(V,E)$ の枝 $e\in E(\partial^+e=u,\partial^-e=v)$ を**図 2** のように 2 本の直列枝 $e',e''(\partial^+e''=u,\partial^-e''=\partial^+e'=w,\partial^-e=v)$ で置き換えることは，計算過程においては，$w:=u$ という計算ステップを v の計算ステップの直前に挿入し，v の計算ステップの右辺にある u を w に置き換えるという自明な修正（**変数の複写と呼ぶ**）に対応する．G に対しいくつかの変数の複写を行ったグラフ $\bar{G}$ を作り，$\bar{G}$ の可縮部分グラフを縮約することが有効である場合も少なくない．

3. 可縮部分グラフの族の性質

本章では，一つの無閉路グラフ $G=(V,E)$ を固定し，その可縮部分グラフの族の性質を扱う．

$$\mathcal{E}=\{E'\subset E\,|\,(V',E')\text{ は }G\text{ の可縮部分グラフ}\}$$

とおく．孤立点をもたない可縮部分グラフ全体と $\mathcal{E}$ は 1 対 1 に対応し，$\phi\in\mathcal{E}$，$E\in\mathcal{E}$ である．E を頂点集合とする無向グラフ (E,R) を，

$$\{e_1,e_2\}\in R\Leftrightarrow[\partial^+e_1=\partial^+e_2\ \text{ または }\ \partial^-e_1=\partial^-e_2]$$

で定義し，その連結成分を $\{C_k\}_k(\phi\neq C_k\subset E)$ と書く．

命題 3.1 $E'\subset E$ に対し，$E'\in\mathcal{E}$ であることは E' がいくつかの C_k の合併の形に書けることと同値である． □

したがって，G の可縮部分グラフへの最も細かい分解は C_k の定める部分グラフの族で与えられる．また，この分解は $O(|E|)$ の手間で構成できる．

命題 3.2 傘は点誘導部分グラフである（すなわち，(V',E') が傘で $e\in E$, $\{\partial^+e,\partial^-e\}\subset V'$ ならば $e\in E'$）． □

$$\mathcal{V}_*=\{V'\subset V\,|\,(V',E')\text{ は }G\text{ の傘}\},$$
$$\mathcal{V}_1=\{V'\subset V\,|\,(V',E')\text{ は }G\text{ の閉傘}\}$$

とおくと，命題 3.2 により，G の傘の全体，閉傘の全体はそれぞれ $\mathcal{V}_*$，$\mathcal{V}_1$ と 1 対 1 に対応する．

命題 3.3 $\mathcal{V}^*$, $\mathcal{V}_1$ はそれぞれ intersecting family を成す（すなわち，$V_1, V_2 \in \mathcal{V}_*$, $V_1 \cap V_2 \neq \phi$ ならば $V_1 \cup V_2, V_1 \cap V_2 \in \mathcal{V}_*$；$\mathcal{V}_1$ についても同様）. □

この命題により，$\mathcal{V}_*$ の極大元は互いに交わらないので，Gのすべての傘を縮約するという操作が矛盾なく定義できる．閉傘についても同様である．

4. 傘と閉傘を求める算法

無閉路グラフ $G=(V, E)$ に新たに2頂点 s，t を導入し，無閉路グラフ $G^*=(V^*, E^*)$ をつくる．ただし，

$$V^* = V \cup \{s, t\},$$

$$E^* = E \cup \{(s, x) \mid x \in X(G)\} \cup \{(f, t) \mid f \in F(G)\}.$$

$u, v \in V^*$ に対して，s から v に至る任意の有向道が u を通るときに $v{\downarrow}u$ とかく．これは s を根とする根付有向木に関する支配関係と呼ばれ，V^* 上の半順序関係である．そのハッセ図 $T{\downarrow}=(V^*, E{\downarrow})$:

$$E{\downarrow} = \{(u, v) \mid v{\downarrow}u;\ u \neq v;\ v{\downarrow}w{\downarrow}u{\downarrow} \text{ ならば } w \in \{u, v\}\}$$

は**支配関係木** (dominator tree) と呼ばれ，$O(|E|\log|E|)$ の手間で構成できる[1)~3)]. 同様にして，$u, v \in V^*$ に対して，u から t に至る任意の有向道が v を通るときに $u{\uparrow}v$ と書き，$T{\uparrow}=(V^*, E{\uparrow})$ を

$$E{\uparrow} = \{(u, v) \mid u{\uparrow}v;\ u \neq v;\ u{\uparrow}w{\uparrow}v \text{ なら } w \in \{u, v\}\}$$

と定義し，さらに $T{\updownarrow}=(V^*, E{\downarrow} \cap E{\uparrow})$ とおく．

命題 4.1 $u, v \in V$ に対し，次の3条件は同値である．

(a) $\{u\}=X(G')$, $\{v\}=F(G')$ である閉傘 G' が存在する．

(b) $v{\downarrow}u$ かつ $u{\uparrow}v$.

(c) $T{\updownarrow}$ において $u \to v$. □

支配関係木 $T{\downarrow}$, $T{\uparrow}$ から $T{\updownarrow}$ を構成し，$T{\updownarrow}$ から頂点 s，t を除去したグラフの連結成分を求めることにより，$O(|E|\log|E|)$ の計算量で，Gの極大な閉傘をすべて列挙することができる．

変数の複写を許すとした場合の極大な傘の全体も以下のように支配関係木 $T{\uparrow}$ を利用して $O(|E|\log|E|)$ の手間で求められる．$T{\uparrow}$ から頂点 t を除去したグラフの連結成分分解を $\{V_k\}_k$（ただし $\phi \neq V_k \subset V$）とする．

$$U_k = \{v \in V_k \mid \partial^+\partial^- v \cap V_k \neq \phi,\ \partial^+\partial^- v - V_k \neq \phi\},$$

$$S_k = \cup \{e \in E \mid \partial^+ e \in V - V_k,\ \partial^- e \in U_k\},\quad S = \cup_k S_k$$

とおく（∂^+, ∂^- はGにおける接続関係を表す）．各 $e \in S$ を変数の複写により2本の直列枝に置き換えたときの枝を e', e''（図2）と表すとして，

$$W_k = \{\partial^+ e' \mid e \in S_k\},\quad W = \cup_k W_k,$$

$$S'_k = \{e' \mid e \in S_k\},\quad S' = \cup_k S'_k,$$

$$S''_k = \{e'' \mid e \in S_k\},\quad S'' = \cup_k S''_k$$

とおく．Gにおいて各 $e \in S$ を変数の複写により2本の直列枝に置き換えて得られるグラフを $\bar{G}=(V \cup W, (E-S) \cup S' \cup S'')$ とする．

命題 4.2 $\bar{G}$ において $V_k \cup W_k$ の誘導する点部分グラフ $\bar{G}_k$ は根付有向木であり，$\{\bar{G}_k\}_k$ がGにおいて変数の複写を許した場合の極大な傘の全体に一致する． □

5. あとがき

計算グラフの上で高速微分法によって偏導関数値を計算すると，関数値に含まれる丸め誤差の推定もできる．傘の縮約によってその計算グラフの構造は変わるが，一つの傘に一つの作業領域を余分に与えれば，同様にして丸め誤差の推定ができる[5)].

参考文献

1) Aho, A. V., Hopcroft, J. E. and Ullman, J. D.: *The Design and Analysis of Computer Algorithms*, Addison-Wesley, Reading, Massachusetts (1974).
2) Aho, A. V., Sethi, R. and Ullman, J. D.: *Compilers—Principles, Techniques, and Tools*, Addison-Wesley, Reading, Massachusetts(1986).
3) Aho, A. V. and Ullman, J. D.: *The Theory of Parsing, Translation, and Compiling, Vol. 2: Compiling*, Prentice-Hall, Englewood Cliffs, N. J. (1973).
4) 伊理正夫，久保田光一：高速微分法とその応用，第7回数理計画シンポジウム論文集，pp. 159-184 (1986).
5) 久保田光一，伊理正夫：高速自動微分法のための処理系（投稿中）.
6) Richardson, D.: Some Undecidable Problems Involving Elementary Functions of a Real Variable, *J. Symbolic Logic*, Vol. 33, pp. 514-520 (1968).
7) 戸川隼人：伊理のアルゴリズムの線形計算向きチューニング，応用数学合同シンポジウム研究報告集，pp. 1-2 (1987).
8) 吉田利信：グラフの変形を用いた偏導関数の計算過程の導出，情報処理学会論文誌，Vol. 28, No. 11, pp. 1112-1120 (1987).

（昭和63年1月11日受付）
（平成元年2月14日採録）

室田 一雄（正会員）

昭和53年東京大学工学部計数工学科卒業．昭和55年同大学院修士課程修了．同年東京大学計数工学科助手．昭和58年筑波大学社会工学系講師．昭和61年東京大学計数工学科助教授，現在に至る．工学博士．数値計算法，グラフ・マトロイド理論などの研究に従事．著書：Systems Analysis by Graphs and Matroids—Structural Solvability and Controllability, Springer, 1987.

久保田光一（正会員）

1960年生．1983年東京大学工学部計数工学科卒業．1985年同大学院修士課程修了．同年東京大学工学部計数工学科助手．数値計算の研究に従事．ACM，日本OR学会各会員．