Gradient と Subgradient

福島　雅夫　京都大学

1. 関数のグラフと勾配

関数の勾配 (gradient) とは文字どおりある点でその関数がどの方向にどのくらい傾いているかを表わすものである．まず微分法の教科書でよく見かける1変数関数のグラフから考えてゆこう(**図1**)．関数$f(x)$が$x=a$で微分可能なとき，点$A=(a, f(a))$を通り微分係数$df(a)/dx$を勾配とする直線

$$y=f(a)+(df(a)/dx)(x-a)$$

を，点Aにおけるfのグラフの**接線**と呼び，その勾配を関数$f(x)$の点aにおける**勾配**と定義する．図1では接線はその名の通り点Aでfのグラフと1点で接しているが，一般には必ずしもこのように“きれい”な関数ばかりではなく，上式で与えられる直線はfのグラフと$x=a$を含むある区間で重なったり，または何度も交差する場合がむしろ普通である．

ここで，fの勾配に対するもう1つの特徴づけを行なっておこう．一般に，直線や平面に垂直なベクトルを**法線**と呼ぶ．図1において，ベクトルABはy成分が-1であるような接線に対する法線ベクトルを表わしている．そのとき，ABのx成分が勾配$df(a)/dx$に等しいことは明らかであろう．このようにして接線に対する法線から勾配を特徴づける考え方は，もっと一般的な勾配を考えるときに役立つ．

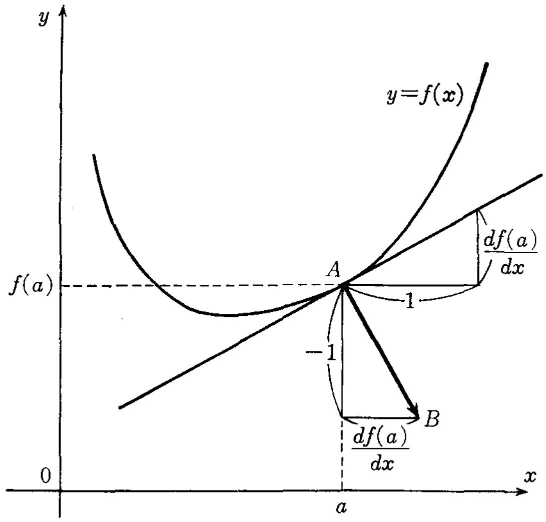

図 1　1変数関数の接線と勾配

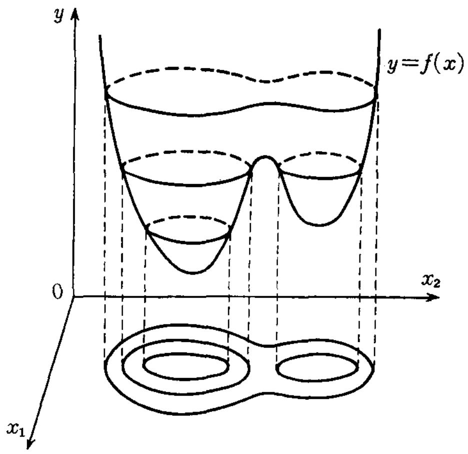

図 2　2変数関数のグラフと等高線

図2の曲面は2変数関数$f(x_1, x_2)$のグラフを立体的に描いたものである．このグラフをいくつかの水平な平面で切った切り口をx_1x_2平面(以下x平面と呼ぶ)に投影したものが関数fの**等高線**である．関数$f(x)$を点$x=a$において(全)微分可能と仮定し，**勾配ベクトル**$\nabla f(a)$を

$$\nabla f(a)=(\partial f(a)/\partial x_1, \partial f(a)/\partial x_2)^T$$

で定義しよう (Tは転置を表わす)．そのとき，点$A=(a, f(a))$におけるfのグラフの**接平面**は

$$\begin{aligned} y&=f(a)+\nabla f(a)^T(x-a)\\ &=f(a)+(\partial f(a)/\partial x_1)(x_1-a_1)\\ &\quad+(\partial f(a)/\partial x_2)(x_2-a_2) \end{aligned}$$

で与えられる．これは関数fのグラフが点Aにおいてx_1方向に$\partial f(a)/\partial x_1$, x_2方向に$\partial f(a)/\partial x_2$だけ傾いていることを示している．たとえば，点$A$を通り$x_1$軸に垂直な(すなわち$x_2y$平面に平行な)平面で$f$のグラフと接平面を切った切り口はちょうど図1のような形になり，その接線の傾きが偏微分係数$\partial f(a)/\partial x_2$に等しい．

それでは，fのグラフに対して勾配ベクトル$\nabla f(a)$は

どのように表わせばよいであろうか．図3を見てみよう．ベクトルABはy成分が-1であるような接平面に対する法線である．1変数の場合と同様に考えれば

$$AB=(\nabla f(a), -1)$$

という関係を得る．すなわちベクトルABをx平面に投影して得られるベクトルが$\nabla f(a)$に他ならない．

図4は図3のx平面を取り出したものである．このように，微分可能な関数に対しては勾配ベクトルは等高線に垂直であり関数値の増加率が最大となる方向を表わしている．また，fの関数値が減少する方向すなわち**降下方向**は勾配ベクトルと鈍角をなすこともわかる．

3変数関数$f(x_1, x_2, x_3)$に対しては最早そのグラフを描くことはできないが，関数fのx空間における**等高面**，すなわちある定数kに対して$f(x_1, x_2\ x_3)=k$となるような点からなる曲面と勾配ベクトルの関係を見ることはできる．**図5**は，関数fが点$x=a$において微分可能ならば，勾配ベクトル

$$\nabla f(a)=(\partial f(a)/\partial x_1, \partial f(a)/\partial x_2, \partial f(a)/\partial x_3)^T$$

は$x=a$を通る等高面（の接平面）に垂直であり，さらにfの降下方向は$\nabla f(a)$と鈍角をなすような方向であることを示している．

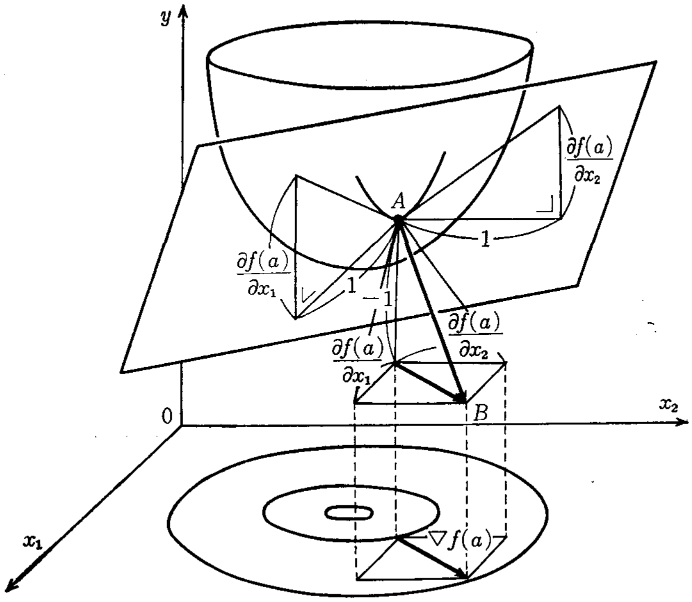

図 3　2変数関数の接平面と勾配

2. 微分不可能関数と劣勾配

上に述べた勾配の概念を微分不可能な関数に対して拡張しようとする試みは，特に最適化理論の分野で近年活発に行なわれている．以下ではそれらの代表的な考え方を簡単に説明しよう．まず，いくつかの用語を定義する．関数$f(x)$に対して，そのグラフより上にある点の集合$\{(x,y) | y \geqq f(x)\}$をfの**エピグラフ**と呼ぶ．集合Sと点$z \in S$に対して

$$p^T(x-z) \leqq 0 \quad \forall x \in S$$

を満たすベクトルpを点zにおけるSの**法線**と呼ぶ（これは上で用いた直線や平面に対する法線の定義の拡張になっている）．

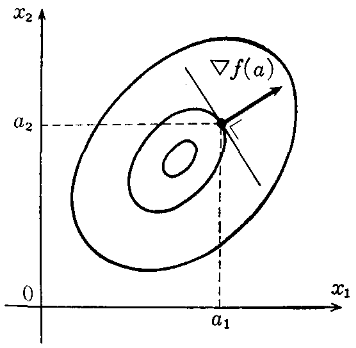

図 4　2変数関数の等高線と勾配ベクトル

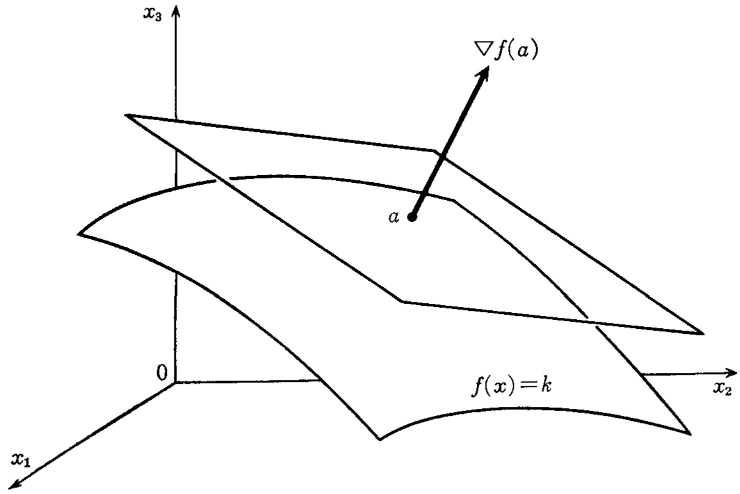

図 5　3変数関数の等高面と勾配ベクトル

1変数関数から始めよう．ここでは以下に述べる多変数関数の場合も含めて，関数は連続かつ各点で任意の方向に関して方向微分係数をもつものとし，さらに微分不可能な点では常にグラフは“下に”尖っていると仮定する．つまり，グラフのカド点は**図6**の点Aのような形をしており，点Cのようなものは含まれないとする．そのとき，任意の点において“左微分係数≦右微分係数”なる関係が成立する．

そこで，点$A=(a,f(a))$を通り傾きが左微分係数と右微分係数の範囲にあるすべての直線を点Aにおけるfの接線とみなし，それらの各直線の勾配をfの勾配（のようなもの）と考えて，関数$f(x)$の$x=a$における**劣勾配**(subgradient)と呼び，その全体の集合を$\partial f(a)$で表わす．

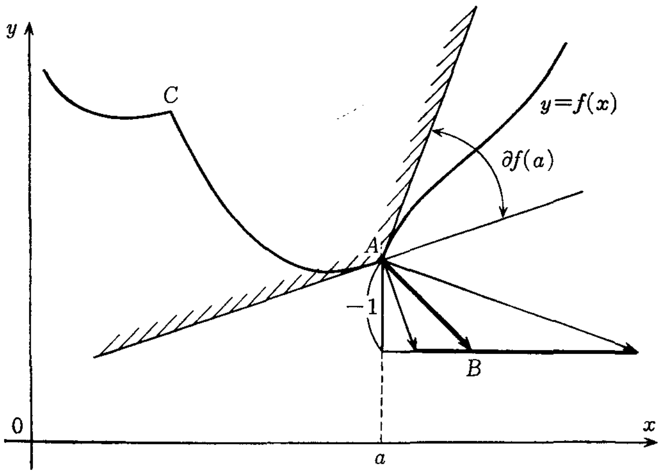

図 6　1変数関数の接錐と劣勾配

次に劣勾配のもうひとつの特徴づけを示しておこう．関数fを点Aにおいて左右両方向を別々に1次近似した関数のエピグラフは点Aを頂点とする凸錐になる（図6の斜線部分）．これを関数fのエピグラフの点Aにおける**接錐**と呼ぶ．この接錐の法線でy成分が－1であるようなものを任意に選べば（図6のベクトルAB），微分可能な場合と同様，そのx成分はfの$x=a$における1つの劣勾配になっている．よって，これを劣勾配の定義とすることも可能である．実際，多変数関数に対して劣勾配の概念を拡張する場合には，むしろこの考え方のほうがわかりやすい．

2変数関数$f(x_1,x_2)$を微分不可能点$x=a$を中心として各方向ごとに1次近似した関数のエピグラフは，**図7**に示すように点$A=(a,f(a))$を頂点とする錐になる．これをfのエピグラフの点Aにおける**接錐**と呼ぶ．ここで1変数の場合と同じように，この接錐の法線でy成分が－1であるような任意のベクトル（たとえば，図7のベクトルAB）のx成分を関数fの$x=a$における劣勾配と定義し，その全体からなる集合を$\partial f(a)$で表わす．集合$\partial f(a)$は常に閉凸集合である．

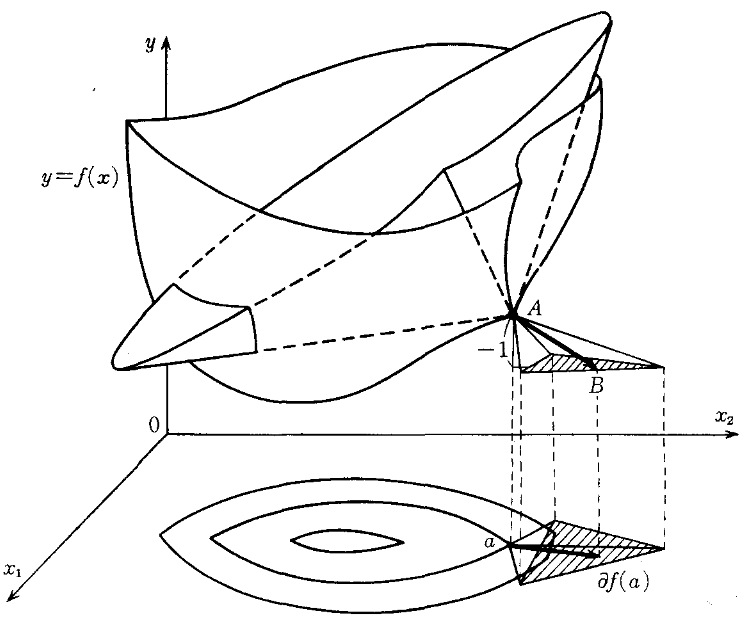

図 7　2変数関数の接錐と劣勾配

図8はx平面でのfの等高線と劣勾配$\partial f(a)$の関係を表わしたものである．微分可能な点$x=a$においては勾配ベクトルの逆方向$-\nabla f(a)$は最急降下方向であったが，微分不可能な点においては劣勾配ベクトルの逆方向は必ずしもfの降下方向とはならないことが分る．しかしながら$\partial f(a)$のなかで大きさが最小であるようなベクトル（図8のベクトルd）に対しては，その逆方向はfの最急降下方向になっていることも確かめられる．

劣勾配という概念はもともと凸関数に対して考えられたものである

が，上で述べた劣勾配は凸関数のそれに対する自然な拡張になっており（非凸関数に対する劣勾配を凸関数に対するものと区別して一般勾配（generalized gradient）と呼ぶこともある），さらに，より一般的な関数に対してもさまざまな劣勾配の拡張がなされている．また，本稿では数式や記号の使用をできるだけ避け，直観的な説明に主眼をおいたため，数学的に厳密さを欠く表現を用いた部分が多い．より詳しい説明は下にあげた文献[1]－[5]などを参照していただきたい．

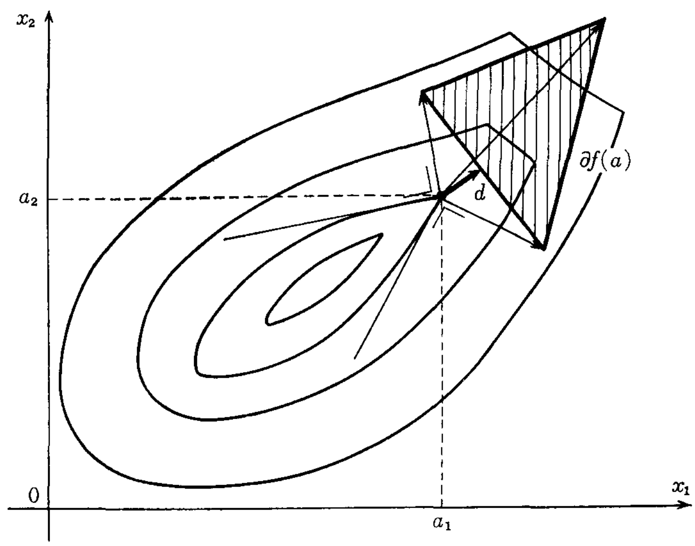

図 8　2変数関数の等高線と劣勾配

参 考 文 献

[1] F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.

[2] R. T. Rockafellar, *The Theory of Subgradients and its Applications to Problems of Optimization*, *Convex and Nonconvex Functions*, Heldermann Verlag, Berlin, 1981.

[3] 今野浩，山下浩，非線形計画法，日科技連，1978.

[4] 志水清孝，相吉英太郎，数理計画法，昭晃堂，1984.

[5] 福島雅夫，非線形最適化の理論，産業図書，1980.