Q. 変数が 3 個(以上)のとき，極値の判定はどのように行えばよいのでしょうか.

A. この質問の答えには 1 年秋学期の「線形代数の応用 1」の知識を必要とする．以下の解説で不明な点はそれらを学習してから考え直してほしい．まず，答えを要約する：

ヘスの行列をつくり，その固有値がすべて正なら極小，すべて負なら極大，正負が混在するなら極値ではない

2 変数の場合，ヘスの行列式は 2 つの固有値の積に一致し，固有値が同符号であることと行列式が正となることが同値である．3 変数以上では，固有値がすべて正ならば行列式が正であるが（その逆は必ずしも成り立たないなど)，行列式の値から固有値の正負を論じることはできない.

●ここでは 3 変数関数$f(x,y,z)$について説明する．3 変数以上への拡張は自明なので省く.

●記号の整理

・変数をベクトルとして扱う：$\vec{r}=(x,y,z)$　およびその微分　$d\vec{r}=(dx,dy,dz)$

・関数：$f(\vec{r})$

・最大傾斜を表すベクトル：$\nabla f=\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}\right)$，および　ナブラ演算子：$\nabla=\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)$

・全微分：$df=\frac{\partial f}{\partial x}dx+\frac{\partial f}{\partial y}dy+\frac{\partial f}{\partial z}dz=\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}\right)\cdot(dx,dy,dz)=\nabla f\cdot d\vec{r}$　(「・」は内積を表す)

●極値の判定条件

・点$\vec{a}$が$f(\vec{r})$極値であるならば，$\vec{a}$における$f(\vec{r})$の接[超]平面が変数の座標[超]平面と平行になる：
$\vec{r}=\vec{a}$において

$$\nabla f(\vec{a})=0 \text{ すなわち } \left.\frac{\partial f}{\partial x}\right|_{\vec{r}=\vec{a}}=0 \text{ かつ } \left.\frac{\partial f}{\partial y}\right|_{\vec{r}=\vec{a}} \text{ かつ } \left.\frac{\partial f}{\partial y}\right|_{\vec{r}=\vec{a}}$$

が成り立つ必要がある．しかし，これらだけでは極値になるか鞍点になるか不明である.

・極値を与える点の候補である$\vec{a}$から$\vec{a}+\vec{h}$まで直線的に移動して関数値の変化を観察する．$\vec{h}$の向きをどのようにとっても関数値が増える（減る）とき$\vec{a}$は極小（極大）の点となる.

・$\vec{a}$から$\vec{a}+\vec{h}$までの移動を媒介変数$t$により表現する：

$$\vec{r}(t)=\vec{a}+\vec{h}\,t,\ \ t:0\to 1$$

1 変数関数$f(t)$を定義する:

$$f(t)=f\left(\vec{a}+\vec{h}t\right)$$

・$f(t)$のマクローリン展開：

$$f(t)=f(0)+\left.\frac{df}{dt}\right|_{t=0}t+\frac{1}{2}\left.\frac{d^2f}{dt^2}\right|_{t=0}t^2+\cdots$$

$\left.\frac{df}{dt}\right|_{t=0}=0$である．実際に，全微分　$df=\nabla f\cdot d\vec{r}=\nabla f\cdot\frac{d\vec{r}}{dt}\,dt$　より

$$\frac{df}{dt}=\nabla f\cdot\frac{d\vec{r}}{dt}\to\nabla f\cdot\vec{h}$$

となり，極値の候補では　$\nabla f(\vec{a})=\nabla f|_{\vec{r}=\vec{a}}=0$より

$$\left.\frac{df}{dt}\right|_{t=0}=\nabla f|_{\vec{r}=\vec{a}}\cdot\vec{h}=0$$

よって，マクローリン展開は

$$f(t)=f(0)+\frac{1}{2}\frac{d^2f}{dt^2}\bigg|_{t=0}t^2+\cdots$$

である．$\frac{d^2f}{dt^2}\Big|_{t=0}$が$\vec{h}$によらず符号を変えないことが，極値となるために必要な条件となる．

・2 階導関数$\frac{d^2f}{dt^2}$：

1 階導関数は

$$\frac{df}{dt}=\nabla f\cdot\frac{d\vec{r}}{dt}=\nabla f\cdot\vec{h}$$

であるが，$\nabla f$が$\vec{r}(t)$の関数であることから，2 階導関数を求めるには上式の$f$のところに$\frac{df}{dt}$を代入し

$$\frac{d^2f}{dt^2}=\frac{d}{dt}\left(\nabla f\cdot\vec{h}\right)=\nabla\left(\nabla f\cdot\vec{h}\right)\cdot\vec{h}$$

成分で表示すると煩雑になるため，$(x,y,z)$を$(x_1,x_2,x_3)$と書きなおし総和記号を用いると

$$\nabla f\cdot\vec{h}=\sum_{j=1,2,3}\frac{\partial f}{\partial x_j}h_j$$

$$\nabla\left(\nabla f\cdot\vec{h}\right)\cdot\vec{h}=\sum_{k=1,2,3}\frac{\partial}{\partial x_k}\left(\nabla f\cdot\vec{h}\right)\cdot h_k=\sum_{j=1,2,3}\frac{\partial}{\partial x_k}\left(\sum_{j=1,2,3}\frac{\partial f}{\partial x_j}h_j\right)\cdot h_k=\sum_{j=1,2,3}\sum_{j=1,2,3}\frac{\partial}{\partial x_k}\frac{\partial f}{\partial x_j}h_jh_k$$

$$=\sum_{j,k=1,2,3}\frac{\partial^2 f}{\partial x_k\partial x_j}h_kh_j$$

を得る．

・$j,k$が 1 と 2 だけを動くとき，2 変数関数の極値判別に用いた式となっていることに注意せよ．

また，$n(>3)$変数のときには$j,k$を 1 から$n$まで動くとすれば，まったく同じ式となっている．

・結論：

$$\frac{d^2f}{dt^2}\bigg|_{t=0}=\nabla\left(\nabla f\cdot\vec{h}\right)\Big|_{\vec{r}=\vec{a}}\cdot\vec{h}=\sum_{j,k=1,2,3}\frac{\partial^2 f}{\partial x_k\partial x_j}\bigg|_{\vec{r}=\vec{a}}h_kh_j$$

が，$\underline{\vec{h}\text{をどのように選んでも}}$符号が変わらないならば，点$\vec{a}$は極値を与える．

以上，2 変数のときとまったく平行に議論が進んだ．

●ヘスの行列

・上の結論の下線部と同等の内容を，$\vec{h}$を用いずに表現したい：

$\frac{\partial^2 f}{\partial x_k\partial x_j}\Big|_{\vec{r}=\vec{a}}$を(j,k)要素とする「ヘスの行列」

$$H=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1\partial x_2} & \frac{\partial^2 f}{\partial x_1\partial x_3}\\ \frac{\partial^2 f}{\partial x_2\partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2\partial x_3}\\ \frac{\partial^2 f}{\partial x_3\partial x_1} & \frac{\partial^2 f}{\partial x_3\partial x_2} & \frac{\partial^2 f}{\partial x_3^2}\end{pmatrix}=\begin{pmatrix}f_{xx} & f_{yx} & f_{zx}\\ f_{yx} & f_{yy} & f_{yz}\\ f_{zx} & f_{zy} & f_{zz}\end{pmatrix}$$

を観察する．

・この行列の特徴は$\frac{\partial^2 f}{\partial x_k\partial x_j}=\frac{\partial^2 f}{\partial x_j\partial x_k}$，すなわち実対称行列である（実数を要素とし転置に対して不変）．

・判定条件の式

$$\left.\frac{d^2 f}{dt^2}\right|_{t=0} = \sum_{j,k=1,2,3} \left.\frac{\partial^2 f}{\partial x_k \partial x_j}\right|_{\vec{r}=\vec{a}} h_k h_j$$

を，ベクトル$\vec{h}$と，$\vec{h}$に$H$を演算した結果であるベクトル $H\vec{h}$の内積，すなわち$\vec{h}\cdot H\vec{h}$と読みなおすことができる．線形代数の記法にしたがうと，$\vec{h}$を縦ベクトル，${}^t\vec{h}$を横ベクトルとして

$$\left.\frac{d^2 f}{dt^2}\right|_{t=0} = \sum_{j,k=1,2,3} \left.\frac{\partial^2 f}{\partial x_k \partial x_j}\right|_{\vec{r}=\vec{a}} h_k h_j = {}^t\vec{h}H\vec{h},$$

である．

・線形代数の定理によると，実対称行列は直交変換$R$（すなわち座標回転）により対角行列$D = RHR^{-1}$となり対角線上には実数の固有値 $\lambda_1, \lambda_2, \lambda_3$が並び，他の要素がすべて0となる:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

・ベクトル$\vec{h}$を$R$により回転したものを$\overrightarrow{h'} = R\vec{h}$とすると

$${}^t\vec{h}H\vec{h} = {}^t\vec{h}R^{-1}\,RHR^{-1}\,R\vec{h} = {}^t\vec{h}R^{-1}\,D\,R\vec{h} = {}^t\!\left(R\vec{h}\right)D\,(R\vec{h}) = {}^t\overrightarrow{h'}D\,\overrightarrow{h'} = h_1'^2\lambda_1 + h_2'^2\lambda_2 + h_3'^2\lambda_3$$

である．

・「どのような$\vec{h}$についても」という条件を「どのような$\overrightarrow{h'} = R\vec{h}$についても」と読み替えると，$\overrightarrow{h'}$によらず

$$\left.\frac{d^2 f}{dt^2}\right|_{t=0} = {}^t\overrightarrow{h'}D\,\overrightarrow{h'} > 0 \ (< 0)$$

となるのはすべての固有値が正（負）のときに限る．正は極小，負は極大のときである．

・2変数のとき，ヘスの行列は

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

である．これを対角化して

$$D = RHR^{-1} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

となったとき，その行列式は

$$\det D = \det RHR^{-1} = \det R \times \det H \times \det R^{-1} = \det R \times \det H \times \frac{1}{\det R} = \det H$$

$$\det H = \det D = \det \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \lambda_1 \lambda_2$$

したがってヘスの行列の行列式が正ならば，$\lambda_1 > 0$かつ$\lambda_2 > 0$となり極小，あるいは$\lambda_1 < 0$かつ$\lambda_2 < 0$となり極大である．

以上