数理計画と最適化アルゴリズム

最適化数学B　252S1527

7月16日（水）

新潟大学理学部数学プログラム
大学院自然科学研究科
　数理物質科学専攻数理科学コース
田 中　　環

数理計画とは

福島雅夫著「数理計画入門」(朝倉書店)から抜粋

1. 最適化すべき問題をいくつかの変数と数式を含む数学モデルに定式化し，それを定められた計算手順(アルゴリズム)を用いて解くための方法論。
2. オペレーションズリサーチの主要なテーマのひとつとして位置づけられ，システム科学，情報科学，経営科学などの分野の基礎的な役割を果たす。
3. 元の問題にかかわらず数学的構造が同じであれば，共通の方法が適用できる，システム的アプローチの典型手法である。
4. 数理計画問題の主な型には，線形計画，非線形計画，ネットワーク計画，組み合わせ計画がある。

線形計画問題の具体例 (Linear Programming Problem)

Maximize $x + y$ **目的関数**

Subject to $3x + 2y \leq 12$

$x + 2y \leq 8$ **制約条件**

$x \geq 0, y \geq 0$

この目的関数は，xとyの関数なので $f(x, y)$ と表せるが，2つの量 xとyを

$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ と1つの文字でベクトルを表すと，上の目的関数は，一般に $f(\mathbf{x})$ と表してよい。

同様にして他の制約関数も次のように表せる。

$$g_1(x, y) = 3x + 2y - 12 \text{とおいて} \Rightarrow g_1(\mathbf{x}) \leq 0$$
$$g_2(x, y) = x + 2y - 8 \quad \text{とおいて} \Rightarrow g_2(\mathbf{x}) \leq 0$$

(定式化された)一般の最適化問題

目的関数 $f(x) \longrightarrow$ **最小化**

制約条件

$$g_1(x) \leqq 0, \cdots, g_m(x) \leqq 0$$
$$h_1(x) = 0, \cdots, h_l(x) = 0$$
$$x \in X$$

$$\Downarrow$$

$$x \in S$$

実行可能領域

Feasible region

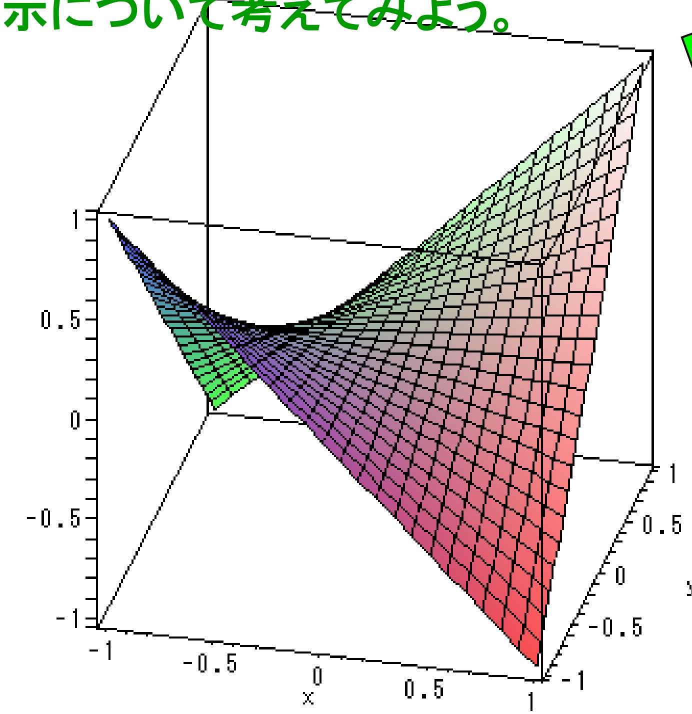
このような問題を数理計画問題といい，解を求める手法を数理計画法という。関数 *f*, *g*, *h* が一次式であるとき線形計画といい，そうでないとき，非線形計画といいます。線形の場合は，有名なシンプレックス法や内点法でおおむね解けています。

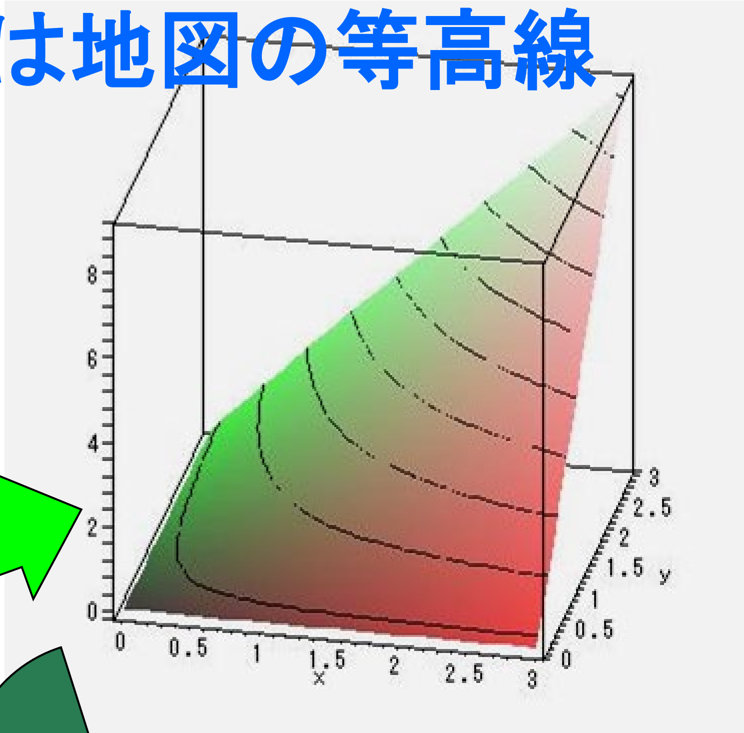
二変数関数の等高線は地図の等高線

式による表現:

関数 $z = xy$

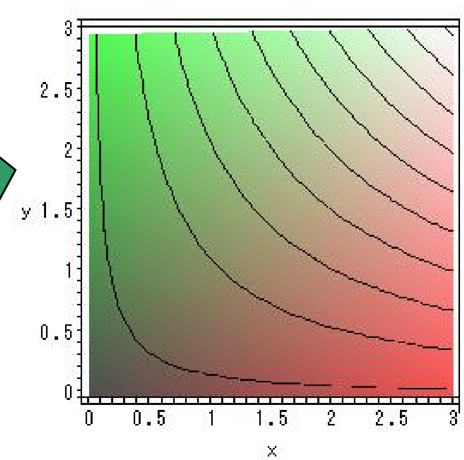
2変数関数のグラフは山や谷のような地形を表しているようにも考えることができる。
社会科の時間によく慣れ親しんだ等高線表示について考えてみよう。





真上から見てみると

$$y = \frac{a}{x}$$



色々な二変数関数の等高線を描いてみよう！

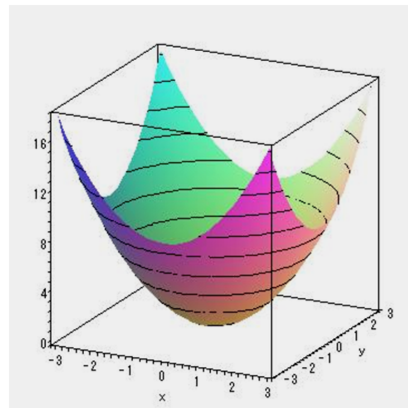
2変数関数のグラフは等高線表示をさせて真上から見てみると意外な**曲線群の集まり**だと分かる。先の例では，反比例を表す分数関数のグラフがたくさん描かれていた。次のような関数はどうだろうか？各自でやってみよう。また，**一方の変数の値をゼロにすることで，グラフの切り口からもとのグラフを想像できはしないだろうか。**

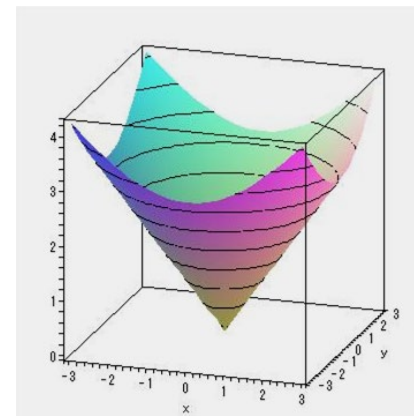
(1) $z = x^2 + y^2$

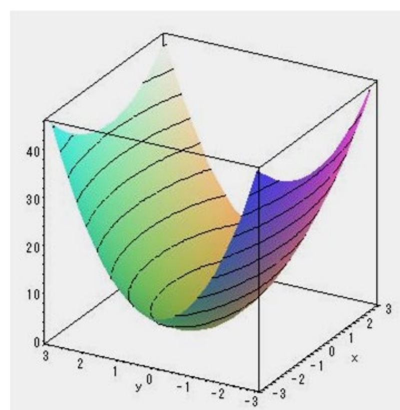
(2) $z = \sqrt{x^2 + y^2}$

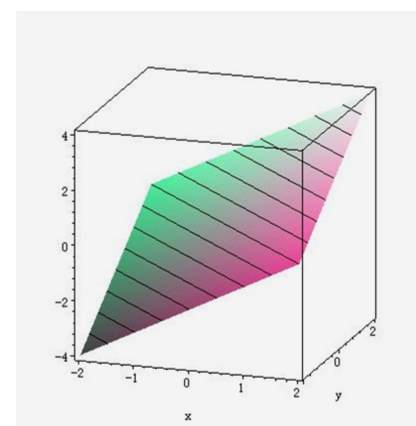
(3) $z = x^2 + 4y^2$

(4) $z = x + y$









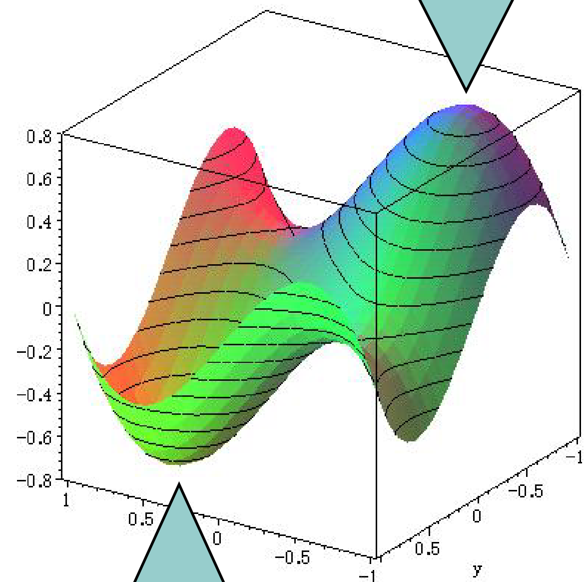
2変数関数の極大･極小

2変数関数 *f(x,y)* が点 (*a,b*) の近くで **<u>丘のようなグラフ</u>**をしている時，つまり，

$f(x,y) \leq f(a,b)$ を満たす時，関数*f(x,y)* は点(*a,b*) で**極大値**をとる，または**極大になる**という。

2変数関数 *f(x,y)* が点 (*a,b*) の近くで **<u>盆地のようなグラフ</u>**をしている時，つまり，

$f(x,y) \geq f(a,b)$ を満たす時，関数*f(x,y)* は点(*a,b*) で**極小値**をとる，または**極小になる**という。



極小値と極大値を総称して<u>極値</u>という。

勾配ベクトル と ヘッセ行列

2変数関数 $f(x,y)$ の<u>偏微分係数を要素</u>とする次のような2次元ベクトル

$$\nabla f(a,b) := \begin{pmatrix} f_x(a,b) \\ f_y(a,b) \end{pmatrix}$$

※記号∇はナブラと読む。

を点 (a,b) における f の<u>**勾配ベクトル (gradient vector)**</u>という。

2変数関数 $f(x,y)$ の勾配ベクトル $\nabla f(a,b)$ をもう一回 x と y で偏微分したもの，つまり f の<u>2次偏微分係数を要素</u>とする次のような2次の<u>**対称行列**</u>

$$\nabla^2 f(a,b) = \begin{pmatrix} f_{xx}(a,b) & f_{yx}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{pmatrix}$$

を点 (a,b) における f の<u>**ヘッセ行列 (Hessian matrix)**</u>という。

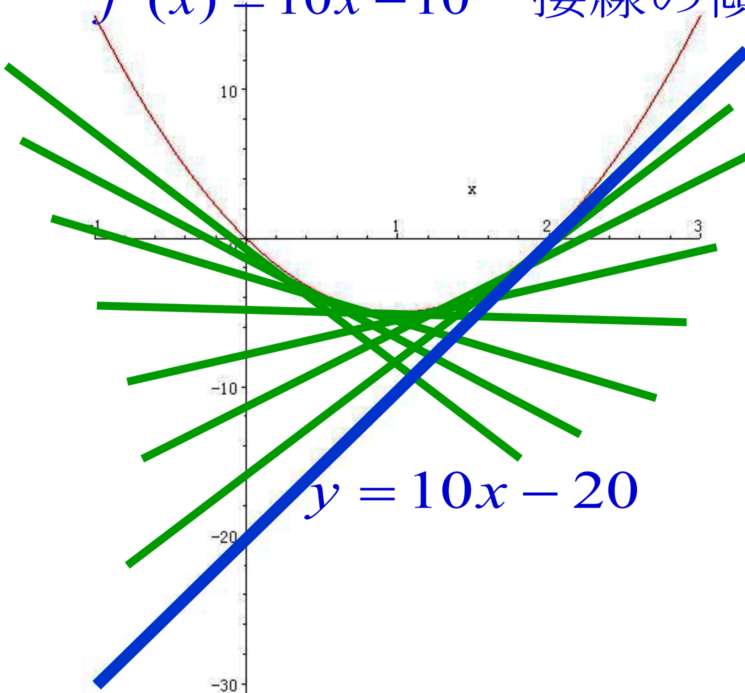
「接線の傾きと微分」と「接平面の傾きと偏微分」

1変数関数のグラフの接線

$$f(x) = 5x^2 - 10x$$

導関数

$f'(x) = 10x - 10$　接線の傾き



$$y = 10x - 20$$

$$y = (10a - 10)(x - a) + f(a)$$

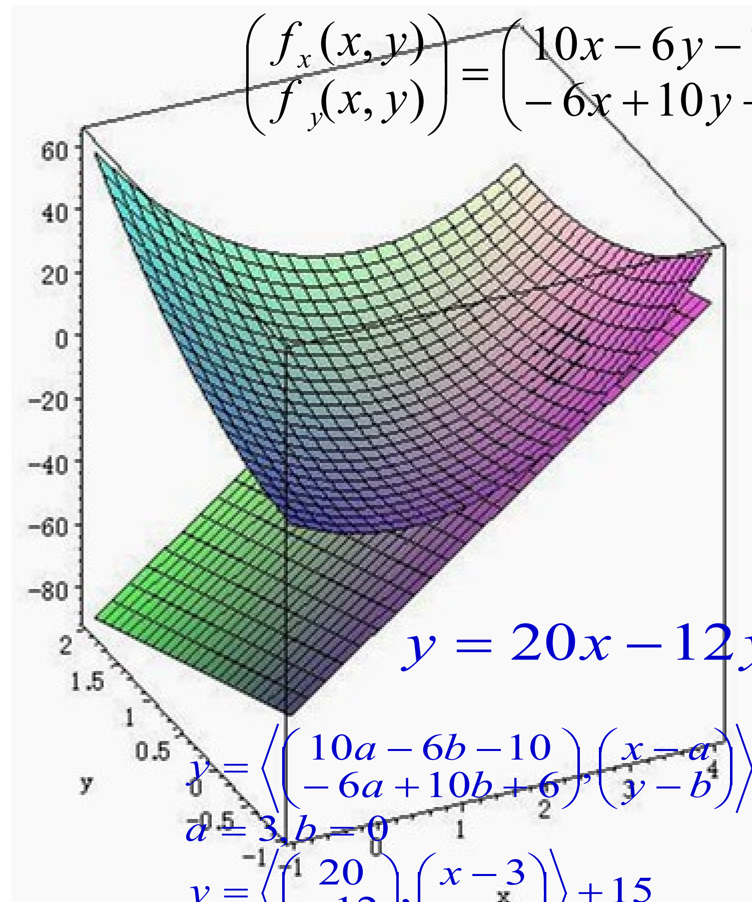
$$a = 2$$

$$y = 10(x - 2)$$

2変数関数のグラフの接線と接平面

$$f(x, y) = 5x^2 - 6xy + 5y^2 - 10x + 6y$$

$$\begin{pmatrix} f_x(x, y) \\ f_y(x, y) \end{pmatrix} = \begin{pmatrix} 10x - 6y - 10 \\ -6x + 10y + 6 \end{pmatrix}$$



$$y = 20x - 12y - 45$$

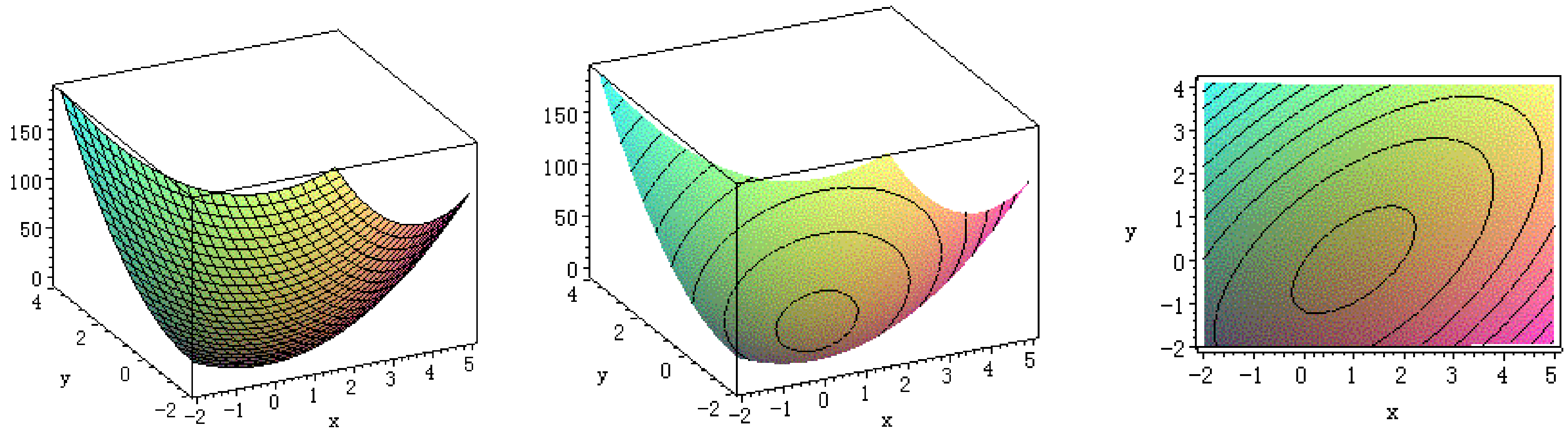
$$y = \left\langle \begin{pmatrix} 10a - 6b - 10 \\ -6a + 10b + 6 \end{pmatrix}, \begin{pmatrix} x - a \\ y - b \end{pmatrix} \right\rangle + f(a, b)$$

$$a = 3,\ b = 0$$

$$y = \left\langle \begin{pmatrix} 20 \\ -12 \end{pmatrix}, \begin{pmatrix} x - 3 \\ y \end{pmatrix} \right\rangle + 15$$

非線形計画問題に対する最適化アルゴリズム

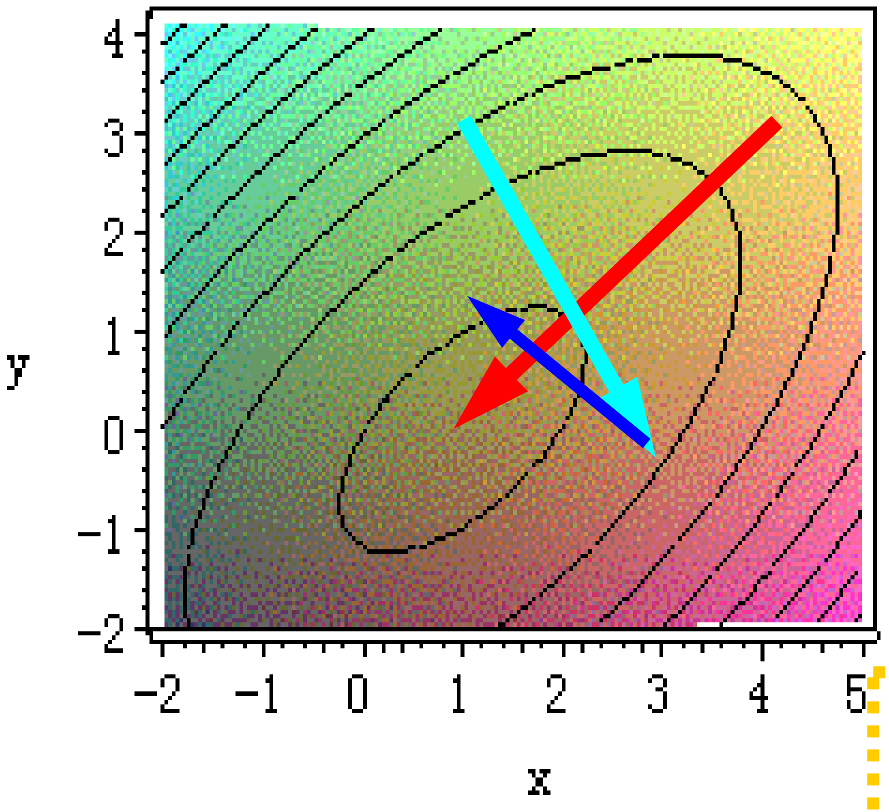
$$f(x, y) = 5x^2 - 6xy + 5y^2 - 10x + 6y$$ **のグラフ**



足から外れたスキー板はその関数を**最小化**するように，スキーヤーの位置から**<u>もっとも急な谷の方向</u>**へ向かって滑ってゆくわけです。

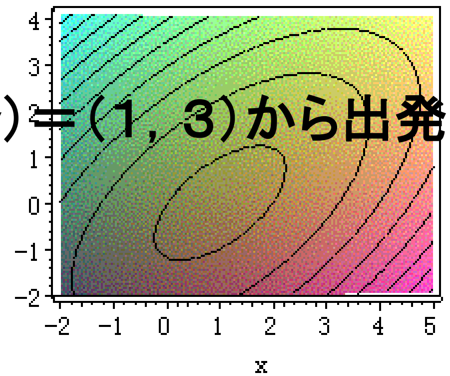
楕円のようなくぼんだところのない図形を凸集合といいます。また，ある高さ以下を現す領域（レベル集合）が凸集合の形の関数を（準）凸関数といいます。今の場合は，連続関数で，凸関数となり，レベル集合が有界閉集合になります。

スキー板が滑ってゆく様子



(1)(x, y)＝(4, 3)から出発する場合

(2)(x, y)＝(1, 3)から出発する場合



降下方向の中で最も急な方向を最急降下方向と呼びます。

勾配ベクトルの計算

$$f(x,y)=5x^2-6xy+5y^2-10x+6y$$

の勾配ベクトルは，各変数について偏微分したものを縦に並べたベクトルになる。**勾配ベクトルとは曲面ののぼりがもっともきつい傾斜を持つ方向を指す。**

$$\nabla f(x,y):=\begin{pmatrix} f_x(x,y) \\ f_y(x,y) \end{pmatrix}=\begin{pmatrix} 10x-6y-10 \\ -6x+10y+6 \end{pmatrix}$$

$$-\nabla f(4,3)=\begin{pmatrix} -12 \\ -12 \end{pmatrix}$$

※降下方向は勾配ベクトルの－1倍，つまり逆方向を指しています。前のグラフでの点(4,3)での降下方向と比較してみよう。点(1,3)ではどうだろうか。

勾配ベクトルをもう一度それぞれ偏微分したものを並べたものがヘッセ行列となる。

$$\nabla^2 f(x,y)=\begin{pmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{pmatrix}=\begin{pmatrix} 10 & -6 \\ -6 & 10 \end{pmatrix}$$

多変数関数の最適性条件

2変数関数の場合

$\mathbf{x}_0 = (x_0, y_0)$において

$$\nabla f(\mathbf{x}_0) = \begin{pmatrix} f_x(\mathbf{x}_0) \\ f_y(\mathbf{x}_0) \end{pmatrix}$$

$$\nabla_2 f(\mathbf{x}_0) = \begin{pmatrix} f_{xx}(\mathbf{x}_0) & f_{xy}(\mathbf{x}_0) \\ f_{yx}(\mathbf{x}_0) & f_{yy}(\mathbf{x}_0) \end{pmatrix}$$

●最適性の1次の必要条件

関数 f が点$\mathbf{x}_0$の近くで連続微分可能とする。

fが点$\mathbf{x}_0$で極値をとる $\Longrightarrow$ $\nabla f(\mathbf{x}_0) = \mathbf{0}$

●最適性の2次の必要条件

関数 f が点$\mathbf{x}_0$の近くで２回連続微分可能とする。

fが点$\mathbf{x}_0$で 極小値（極大値） をとる $\Longrightarrow$
- $\nabla f(\mathbf{x}_0) = 0$かつ
- $\nabla_2 f(\mathbf{x}_0)$が 半正定値（半負定値）

●最適性の2次の十分条件

関数 f が点$\mathbf{x}_0$の近くで２回連続微分可能とする。

$\nabla f(\mathbf{x}_0) = 0$かつ

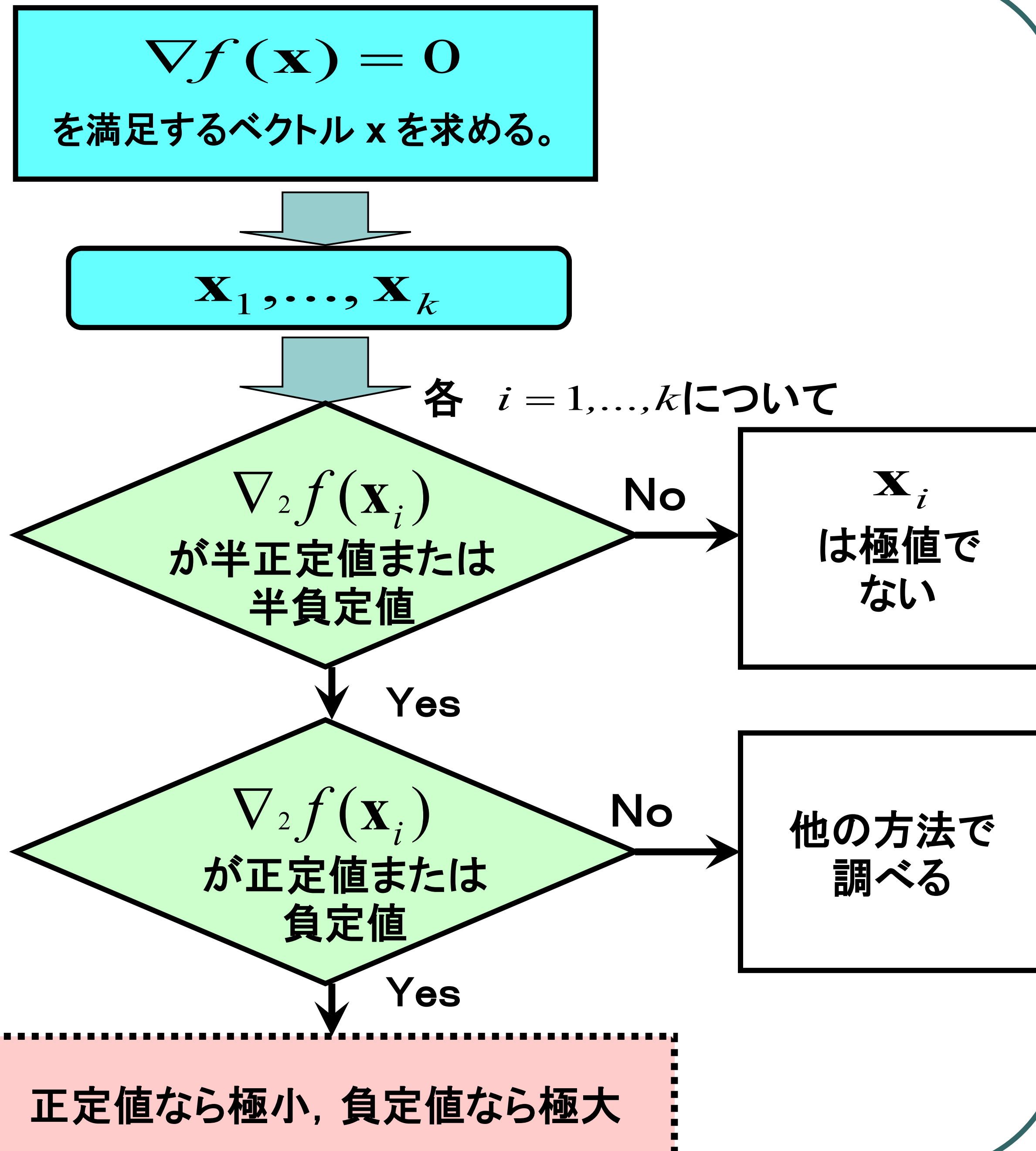
$\nabla_2 f(\mathbf{x}_0)$が 正定値（負定値） $\Longrightarrow$ fが点$\mathbf{x}_0$で 極小値（極大値） をとる

多変数関数の極値判定の仕方

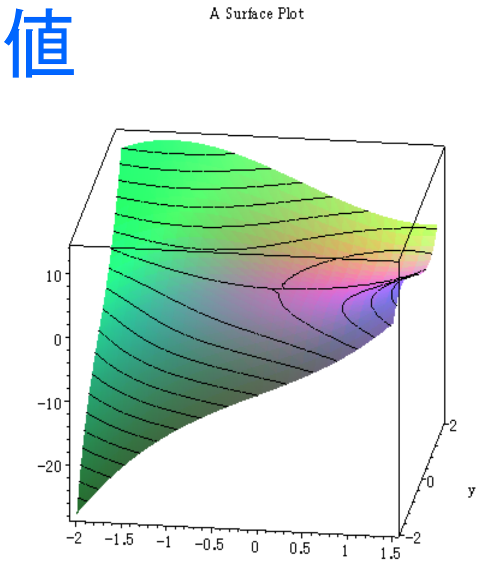
（1）最適性の1次の必要条件に従って極値の候補を探す。（勾配ベクトルがゼロベクトルとなる点）

（2）最適性の2次の必要条件と2次の十分条件に従って極値の候補を絞り込む。（ヘッセ行列の定値性を調べる）

（3）それでも判定不可能なものについては，2方向以上異なる近づき方をしたときに増減の違いがあれば極値とならない。



2変数実数値関数の極値



2変数実数値関数 $f(x,y)=x^3-3xy+y^3$ の極値を調

(1) 最適性の1次の必要条件 ： $\nabla f(x,y)=0$

$$\nabla f(x,y)=\begin{pmatrix}3x^2-3y\\-3x+3y^2\end{pmatrix}$$ より、必要条件を満たす点

$$\begin{cases}f_x(x,y)=3x^2-3y=0\cdots①\\f_y(x,y)=-3x+3y^2=0\cdots②\end{cases}$$

①、②より極値をとる候補の点は $(0,0)$ と $(1,1)$ である。

(2) 最適性の2次の十分条件及び必要条件 ： $\nabla^2 f(x,y)$ の定値性

$$\nabla^2 f(x,y)=\begin{pmatrix}6x&-3\\-3&6y\end{pmatrix}$$ より、(1)で見つけた2点を考えると

行列の定値性を判定する方法は，主に3通りある。固有値判定，小行列式による判定，直接判定。

(i) 点 $(0,0)$ では $\nabla^2 f(0,0)=\begin{pmatrix}0&-3\\-3&0\end{pmatrix}$ なので<u>**半正定値でも半負定値でもない**</u>。

よって，2次の必要条件を満たさないので、極値をとらないことがわかる。

(ii)点 $(1,1)$ では $\nabla^2 f(1,1)=\begin{pmatrix}6&-3\\-3&6\end{pmatrix}$ なので正定値になっており、

2次の十分条件を満たしているので、極小値 $f(1,1)=-1$ をとることがわかる。

小行列式による行列の定値性判定法

行列

$$A=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \qquad A_k=\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

に対して

*A*が正定値であるための必要十分条件は，すべての主対角行列式が正 $\boxed{\det(A_k)>0 \ \ (1\le k\le n)}$ が成立すること。

従って，ヘッセ行列の場合 $\nabla^2 f(a,b)=\begin{pmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{pmatrix}$ に適用すると，

$$\left|f_{xx}(a,b)\right| = f_{xx}(a,b)>0 \qquad \begin{vmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{vmatrix} >0$$

つまり，

$\nabla^2 f(a,b)$ が正定値であるための必要十分条件は，

$$f_{xx}(a,b)>0 \quad \text{かつ} \quad f_{xx}(a,b)f_{yy}(a,b)-f_{xy}(a,b)f_{yx}(a,b)>0$$

前の例で確かめてみよう！

前の例では，２変数実数値関数が次のようになっていたので，

$$f(x,y)=x^3-3xy+y^3$$

$$\nabla f(a,b)=\begin{pmatrix} f_x(a,b) \\ f_y(a,b) \end{pmatrix}$$

$$\nabla f(x,y)=$$

$$\nabla^2 f(a,b)=\begin{pmatrix} f_{xx}(a,b) & f_{yx}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{pmatrix}$$

$$\nabla^2 f(x,y)=$$

$$\nabla^2 f(a,b)=\begin{pmatrix} f_{xx}(a,b) & f_{yx}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{pmatrix}$$

一次の最適性条件より，次の2つの候補

$(x,y)=(0,0)$ $(x,y)=(1,1)$

$$\begin{vmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{vmatrix}=36ab-9$$

について考えてみると，(0,0) は二次の必要条件から解にはなりません。(1,1) は二次の十分条件から最適解になります。