第1章　数学的準備

1.1　多変数関数

$f(x,y)=x^2+y^2$ を使って，2 変数関数の性質を勉強する．まず，いくつかの (x,y) において関数がどのような値をとるか調べると，以下のようになる．

$x \setminus y$	-2	-1	0	1	2
-2	8	5	4	5	8
-1	5	2	1	2	5
0	4	1	0	1	4
1	5	2	1	2	5
2	8	5	4	5	8

次に，関数の値を幾何的にとらえるために $z=x^2+y^2$ とおいて，この関数のグラフを xyz 空間に描いてみよう．

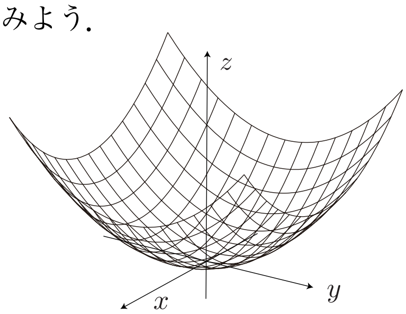

図 1.1: $z=f(x,y)$ のグラフ

等高線

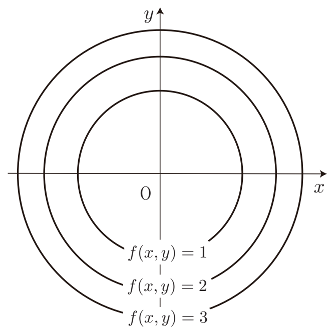

図 1.2: $z = x^2 + y^2$ の等高線

参考に $f(x,y) = -x^3 - 3xy^2 + y^3 + 3x$ のグラフと等高線も挙げておこう．

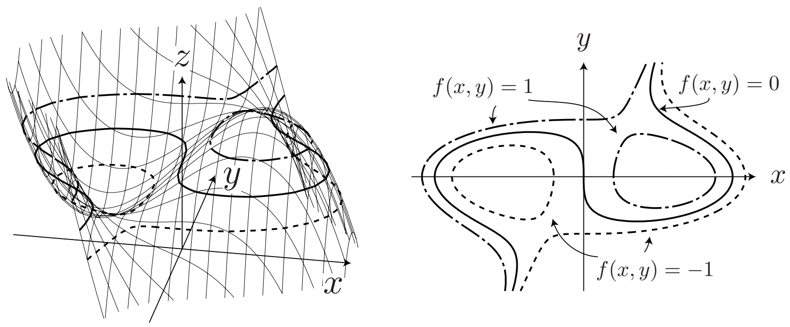

図 1.3: $f(x,y) = -x^3 - 3xy^2 + y^3 + 3x$ のグラフと等高線

1.2 凸関数の性質

1.2.1 凸関数の定義

［定義］ **1.1.** f を n 変数関数とする．$0 < \lambda < 1$ を満たすすべての λ とすべての $u, v \in \mathbb{R}^n$ に対して，

$$f((1-\lambda)u + \lambda v) \leq (1-\lambda)f(u) + \lambda f(v) \tag{1.1}$$

が成り立つとき，f を 凸関数 と呼ぶ．また，

$$f((1-\lambda)u + \lambda v) < (1-\lambda)f(u) + \lambda f(v) \quad (u \neq v) \tag{1.2}$$

が成り立つとき f を 狭義凸関数 と呼ぶ．

$-f$ が凸関数のとき，f を 凹関数 と呼ぶ．狭義凹関数も同様である．

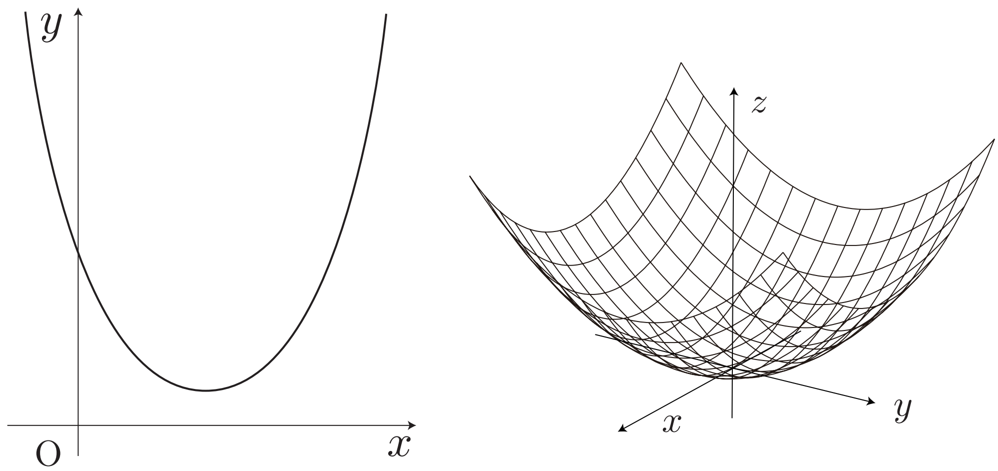

図 1.4: 凸関数の例 左が 1 変数凸関数，右が 2 変数凸関数

[解説]　説明がやさしい順に，はじめに狭義凸関数について解説する．

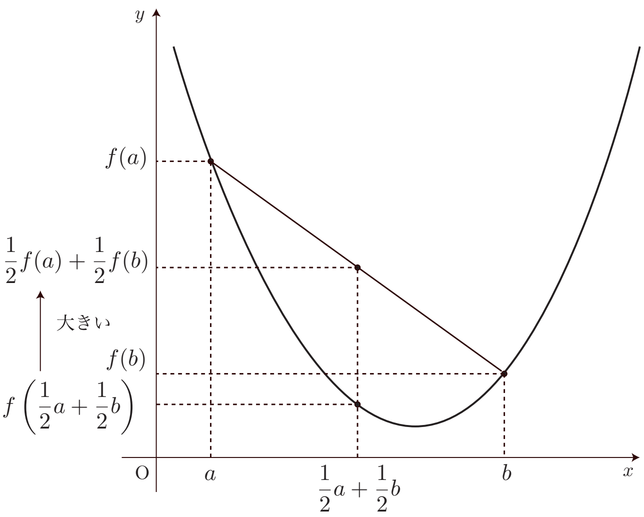

図 1.5: 狭義凸関数に対して，定義式で $\lambda = \frac{1}{2}$ とした場合

式 (1.2) は，

「f が狭義凸関数」

$\Updownarrow$

「点 $(u, f(u))$ と $(v, f(v))$ を結んだ線分 が，$\underline{f \text{ グラフより上部}}$ にある」

ということを表す．

次に，凸関数について考えよう．狭義凸関数と違い式 (1.1) では等号も許されている．これは

「f が凸関数」

$\Updownarrow$

「点 $(u, f(u))$ と $(v, f(v))$ を結んだ線分が，<u>f のグラフ上，またはグラフより上部</u>にある」

ということを表している．

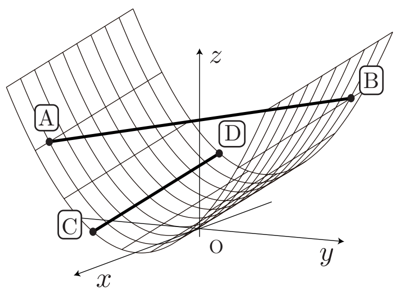

図 1.6: 狭義凸関数でない凸関数の例

1.2.2　凸関数と接平面

2 変数関数 f の点 $(a,b,f(a,b))$ における接平面は，

$$z=f(a,b)+f_x(a,b)(x-a)+f_y(a,b)(y-b)$$

2 変数関数 f が凸ならば，

$$f(x,y)\geq f(a,b)+f_x(a,b)(x-a)+f_y(a,b)(y-b) \tag{1.3}$$

が成り立つ．

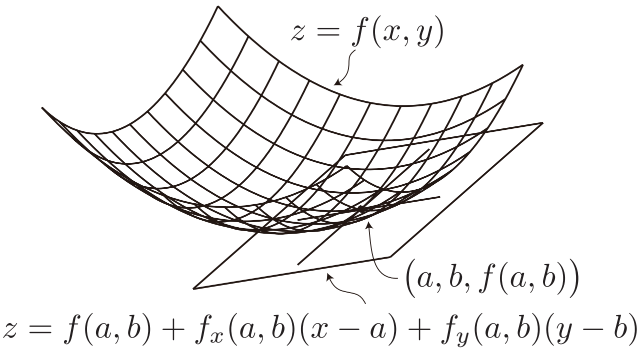

図 1.7: 凸関数と接平面

これは一般に n 変数関数に対しても成り立つので，その証明を以下に挙げる．まず記号を導入する．

勾配ベクトル

［定義］ **1.2.** f を 2 変数関数 f とする．

$$\text{勾配ベクトル}\quad \nabla f(x,y)=\begin{bmatrix} f_x(x,y)\\ f_y(x,y)\end{bmatrix}$$

n 変数関数 f の場合

$$\nabla f(u)=\begin{bmatrix} f_{x_1}(u)\\ \vdots \\ f_{x_n}(u)\end{bmatrix}$$

［例］ **1.3.** $f(x,y)=x^2+3y^2$ とすると，点 $(5,3)$ における勾配ベクトルは，$f_x=2x$, $f_y=6y$ より，

$$\nabla f(5,3)=\begin{bmatrix} f_x(5,3)\\ f_y(5,3)\end{bmatrix}=\begin{bmatrix} 2\cdot 5\\ 6\cdot 3\end{bmatrix}=\begin{bmatrix} 10\\ 18\end{bmatrix}$$

となる．

凸関数と接平面に関する不等式

［命題］ **1.4.** n 変数関数 f に対して，以下が成り立つ：

$$\begin{aligned} & f \text{ が凸関数} \\ \Longleftrightarrow \ & f(v) \geq f(u) + \nabla f(u) \cdot (v-u) \ (\text{すべての } u, v \in \mathbb{R}^n) \end{aligned} \tag{1.4}$$

証明. まず f を凸関数，$u, v \in \mathbb{R}^n$, $0 < \lambda < 1$ とする．凸関数の定義式とテイラーの定理より，

$$\begin{aligned} (1-\lambda)f(u) + \lambda f(v) \geq f\left((1-\lambda)u + \lambda v\right) &= f\left(u + \lambda(v-u)\right) \\ &= f(u) + \lambda \nabla f(u) \cdot (v-u) + o(\lambda \|v-u\|), \\ \lambda f(v) &\geq \lambda f(u) + \lambda \nabla f(u) \cdot (v-u) + o(\lambda \|v-u\|), \\ f(v) &\geq f(u) + \nabla f(u) \cdot (v-u) + \frac{o(\lambda \|v-u\|)}{\lambda} \end{aligned}$$

を得る．$\lambda \to 0$ とすると，$\frac{o(\lambda \|u-v\|)}{\lambda} \to 0$ となるので，不等式 (1.4) が成り立つ．逆に，(1.4) が成り立つとする．このとき，

$$\begin{aligned} f(u) &\geq f((1-\lambda)u + \lambda v) + \lambda \nabla f((1-\lambda)u + \lambda v) \cdot (u-v) \\ f(v) &\geq f((1-\lambda)u + \lambda v) - (1-\lambda) \nabla f((1-\lambda)u + \lambda v) \cdot (u-v) \end{aligned}$$

を得る．ここで，両辺にそれぞれ $(1-\lambda)$ と λ をかけて足すと

$$\begin{aligned} (1-\lambda)f(u) + \lambda f(v) &\geq (1-\lambda)f((1-\lambda)u + \lambda v) + \lambda f((1-\lambda)u + \lambda v) \\ &= f((1-\lambda)u + \lambda v) \end{aligned}$$

が成り立つ．よって定義より f は凸関数である． □

1.3　凸関数の判定

1.3.1　凸性と微分

関数 $f(x) = x^2$ はグラフより，凸関数．それでは，次の関数は？

$$f(x) = \sqrt{1+x^2}$$

［定理］ **1.5.** 1 変数関数 h に対して，以下が成り立つ：

$$h \text{ が凸関数} \Leftrightarrow \text{すべての } t \in \mathbb{R} \text{ に対して } h''(t) \geq 0$$

定理 1.5 を用いると，$f'(x) = \dfrac{x}{\sqrt{1+x^2}}$，$f''(x) = \dfrac{1}{(1+x^2)^{\frac{3}{2}}}$ より，すべての x に対して，$f''(x) > 0$ となるので，$f(x)$ は凸関数である．

ヘッセ行列

次に，2 変数関数

$$g(x,y) = x^2 - 2xy + 2y^2$$

は凸関数？

［定義］ **1.6.** f を 2 変数関数とする．

$$\text{ヘッセ行列 } \nabla^2 f(x,y) = \begin{bmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{bmatrix}$$

講義では，扱う関数は十分滑らかなので，常に $f_{xy} = f_{yx}$ が成り立つ

$f(x,y) = x^3 + 2xy + 3y^2$ とすると，

$$\text{勾配ベクトル } \nabla f(x,y) = \begin{bmatrix} 3x^2 + 2y \\ 2x + 6y \end{bmatrix}, \quad \text{ヘッセ行列は } \nabla^2 f(x,y) = \begin{bmatrix} 6x & 2 \\ 2 & 6 \end{bmatrix}$$

一般の n 変数関数 f と $u = (x_1, \ldots, x_n)$ に対しては，

$$\text{ヘッセ行列 } \nabla^2 f(u) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(u) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(u) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(u) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(u) & & \ddots & \vdots \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(u) & & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(u) \end{bmatrix}$$

ヘッセ行列による条件

　ヘッセ行列を用いると以下のような凸性の判定方法がある．

［定理］ **1.7.** (1). f が凸関数 $\Longleftrightarrow$ 各点 $a \in \mathbb{R}^n$ で，

$$ {}^t u \nabla^2 f(a) u \geq 0 \ (\text{すべてのベクトル } u \in \mathbb{R}^n) $$

が成り立つ．

(2). f が狭義凸関数 $\Longleftarrow$ 各点 $a \in \mathbb{R}^n$ で，

$$ {}^t u \nabla^2 f(a) u > 0 \ (u \neq 0 \text{ となるすべてのベクトル } u \in \mathbb{R}^n) $$

が成り立つ．

証明の概要．$b, c \in \mathbb{R}^n$ に対して，$h(t) = f\left((1-t)b + tc\right)$ とおく．すると

$$ f \text{ が凸} \Longleftrightarrow \text{すべての } b, c \in \mathbb{R}^n \text{ に対して } h(t) \text{ が凸} $$

が成り立つ．また，h は 1 変数関数なので，

$$ h \text{ が凸} \Longleftrightarrow \text{すべての } t \text{ に対して } \tfrac{d^2}{dt^2} h(t) \geq 0 $$

が成り立つ．ここで，

$$ \begin{cases} a := (1-t)b + tc = b + t(c-b) \\ u := c - b \end{cases} $$

とおくと，

$$ \begin{aligned} \frac{d}{dt} h(t) &= \sum_{i=1}^{n} f_{x_i}(a)(c_i - b_i) = {}^t u \nabla f(a), \\ \frac{d^2}{dt^2} h(t) &= \sum_{j=1}^{n} \sum_{i=1}^{n} f_{x_i x_j}(a)(c_i - b_i)(c_j - b_j) = \sum_{j=1}^{n} \left({}^t u \begin{bmatrix} f_{x_1 x_j}(a) \\ \vdots \\ f_{x_n x_j}(a) \end{bmatrix} \right) (c_j - b_j) \\ &= {}^t u \nabla^2 f(a) u \end{aligned} $$

となる．これより (1) が導かれる．(2) も同様である． □

［例］ **1.8.** $f(x,y)=x^2-2xy+2y^2$ とすると，勾配ベクトルとヘッセ行列は

$$\nabla f(x,y)=\begin{bmatrix}2x-2y\\-2x+4y\end{bmatrix},\quad \nabla^2 f(x,y)=\begin{bmatrix}2&-2\\-2&4\end{bmatrix}$$

となる．定理 1.7 を用いて f が凸関数かどうか調べよう．$u={}^t(u_1,u_2)$ とすると，各点 (x,y) で，

$${}^tu\nabla f(x,y)u=[\ u_1\ \ u_2\ ]\begin{bmatrix}2&-2\\-2&4\end{bmatrix}\begin{bmatrix}u_1\\u_2\end{bmatrix}=2u_1^2-4u_1u_2+4u_2^2$$

となる．さらに，すべてのベクトル u に対して

$$2u_1^2-4u_1u_2+4u_2^2=2(u_1^2-2u_1u_2+2u_2^2)=2\left\{(u_1-u_2)^2+u_2^2\right\}\geq 0$$

となるので，定理 1.7 より f は凸関数である．

1.3.2　行列の正値性

2 次形式

［定義］ **1.9.** 行列 A とベクトル $u=(u_1,u_2,\ldots,u_n)$ に対して，$u_1,\ldots,u_n$ を変数に持つ多項式

$$p(u)={}^tuAu$$

を 2 次形式 と呼ぶ．

［例］ **1.10.** 行列 $A=\begin{bmatrix}2&0\\0&1\end{bmatrix}$, に対して $u=(x,y)$ とすると，

$${}^tuAu=[\ x\ \ y\ ]\begin{bmatrix}2&0\\0&1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=[\ x\ \ y\ ]\begin{bmatrix}2x\\y\end{bmatrix}=2x^2+y^2$$

は 2 次形式である．また，$B=\begin{bmatrix}1&1\\1&-1\end{bmatrix}$ とすると，

$$\begin{aligned}{}^tuBu=[\ x\ \ y\ ]\begin{bmatrix}1&1\\1&-1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}=[\ x\ \ y\ ]\begin{bmatrix}x+y\\x-y\end{bmatrix}\\=x^2+xy+yx-y^2=x^2+2xy-y^2\end{aligned}$$

も 2 次形式である．

正値性の定義

［定義］ **1.11.** 行列 A を持つ対称行列 とする．

- すべての $u \in \mathbb{R}^n$ に対して，2 次形式が
$$ {}^t uAu \geq 0 $$
を満たすとき，A を 半正定値 という．また，逆の不等号が成り立つとき，A を 半負定値 という．

- $u \neq 0$ を満たすすべての $u \in \mathbb{R}^2$ に対して，
$$ {}^t uAu > 0 $$
を満たすとき，A を 正定値 という．逆の不等号が成り立つとき，A を 負定値 という．

- $u \in \mathbb{R}^2$ によって ${}^t uAu$ の符号が正にも負にもなるとき，A を不定という．

［例］ **1.12.** 例 1.10 の行列 $A = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ の正値性を調べてみよう．$u = (x, y)$ とすると，2 次形式は ${}^t uAv = 2x^2 + y^2$ となる．ここで，$2x^2 + y^2 > 0$ $(u \neq 0)$ となるので，A は正定値である．

一方，行列 $B = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ に対して，2 次形式は ${}^t uBu = x^2 + 2xy - y^2$ となる．ここで，$u = (1, 0)$ とすると ${}^t uBu = 1 > 0$ となるが，$u = (0, 1)$ とすると ${}^t uBu = -1 < 0$ となるので，B は不定である．

凸関数の判定定理の再掲（正値性を用いて）

これらの言葉を用いると，定理 *1.7* は以下のように書ける .

［定理］ **1.13** (定理 1.7 の言い換え). 関数 f に対して以下が成り立つ：

- f が凸関数
$\Longleftrightarrow$ 各点 $a \in \mathbb{R}^n$ において，ヘッセ行列 $\nabla^2 f(a)$ が半正定値である．

- f が狭義凸関数
$\Longleftarrow$ 各点 $a \in \mathbb{R}^n$ において，ヘッセ行列 $\nabla^2 f(a)$ が正定値である．

固有値を用いた判定法

［定理］ **1.14.** A を対称行列とする．以下の主張が成り立つ．

(1). A が半正定値 $\Longleftrightarrow$ A の固有値 がすべて 0 以上

(2). A が正定値 $\Longleftrightarrow$ A の固有値がすべて正

(3). A が不定 $\Longleftrightarrow$ A が正と負の固有値を持つ

証明. A を $n \times n$ 行列とする．A は対称行列なので，定理 **??** よりある直交行列 P が存在して，

$$P^{-1}AP = D$$

と対角化できる．ここで, D は A の固有値を対角要素に持つ対角行列である．いま，$P^{-1} = {}^tP$ に注意して，$v = {}^tPu$ と変数変換すると， $u = Pv$ より，

$${}^tuAu = {}^t(Pv)A(Pv) = {}^tv({}^tPAP)v = {}^tvDv = \lambda_1 v_1^2 + \cdots + \lambda_n v_n^2$$

を得る．ここで，$v = (v_1, \ldots, v_n)$ であり，$\lambda_1, \ldots, \lambda_n$ は A の固有値である．よって，固有値がすべて 0 以上ならば，${}^tuAu \geq 0$ （すべての u）となり，これより (1) が導かれる．他も同様である． $\square$

［例］ **1.15.** $f(x, y) = 3x^2 - 2xy + 3y^2$ とすると，

$$\nabla f(x, y) = \begin{bmatrix} 6x - 2y \\ 2x - 6y \end{bmatrix}, \quad \nabla^2 f(x, y) = \begin{bmatrix} 6 & -2 \\ -2 & 6 \end{bmatrix}$$

となる．ヘッセ行列は定数行列であり，この固有値は $4, 8$ になる．よって，ヘッセ行列は正定値で，f は狭義凸関数である．

行列式を用いた判定法

［定理］ **1.16.** A を 2×2 行列とする．

(1). $|A| > 0$ かつ $a > 0 \Leftrightarrow A$ は正定値

(2). $|A| > 0$ かつ $a < 0 \Leftrightarrow A$ は負定値

(3). $|A| < 0 \Leftrightarrow A$ は不定

証明. (1) まず，$|A| > 0$ かつ $a > 0$ とする．任意の $(x, y) \in \mathbb{R}^2$ に対して，

$$\begin{aligned} [\, x \quad y \,] A \begin{bmatrix} x \\ y \end{bmatrix} &= ax^2 + 2bxy + dy^2 = a\left(x + \frac{by}{a}\right)^2 + \frac{y^2}{a}(ad - b^2) \\ &= a\left(x + \frac{by}{a}\right)^2 + \frac{1}{a}|A|\ y^2 \end{aligned} \tag{1.5}$$

を得る．ただし，$|A|$ は A の行列式 $ad-b^2$ を表す．いま，$|A|>0,\ a>0$ より

$$a\left(x+\frac{by}{a}\right)^2+\frac{1}{a}|A|\ y^2 \geq 0$$

が成り立つ．ここで $a\left(x+\dfrac{by}{a}\right)^2+\dfrac{1}{a}|A|\ y^2=0$ となるときは，各項が 0 以上なので，$x+\dfrac{by}{a}=0,$ かつ $y=0$ より，$(x,y)=(0,0)$ である．よって (1.5) より，2 次形式は $(x,y)\neq(0,0)$ で正の値をとる．逆に，

$$(x,y)\neq(0,0) \Rightarrow [\ x \quad y\ ]A\begin{bmatrix} x \\ y \end{bmatrix}>0$$

が成り立つとする．この 2 次形式に $(x,y)=(1,0)$ を代入すると，$a>0$ を得る．さらに，定理 **??** より，行列式の値は固有値の積と等しいので，定理 1.14 より $|A|>0$ となる．(2) も同様に示される．

(3) 行列式の値は固有値の積と等しいことから，2×2 行列の場合，行列式が負ということは二つある固有値の符号が異なるということである．よって定理 1.14 より示される． □

［例］**1.17.** 例 1.15 の関数 $f(x,y)=3x^2-2xy+3y^2$ のヘッセ行列 $\nabla^2 f(x,y)=\begin{bmatrix} 6 & 2 \\ 2 & 6 \end{bmatrix}$ の正値性を行列式を用いて調べると，

$$|\nabla^2 f(x,y)|=32>0 \text{ かつ } f_{xx}(x,y)=6>0$$

なので，ヘッセ行列は正定値である．

［練習問題］ **1.** 以下の関数の勾配ベクトルとヘッセ行列を求め，凸関数かどうか調べよ．

(1). $f(x,y)=x^2+xy+y^2$　(2). $f(x,y)=2x^2+4xy+y^2$
(3). $f(x,y)=x^2+xy+y^3$　(4). $f(x,y)=\sqrt{1+x^2+y^2}$
(5). $f(x,y,z)=x^2+y^2+z^2-xy-yz-zx$