確率と最適化

凸関数とその性質

凸計画問題の標準形

$f_0, f_1, \ldots, f_m$ は凸関数。
$x \in \mathbb{R}^n, a_i \in \mathbb{R}^n, b_i \in \mathbb{R}$.

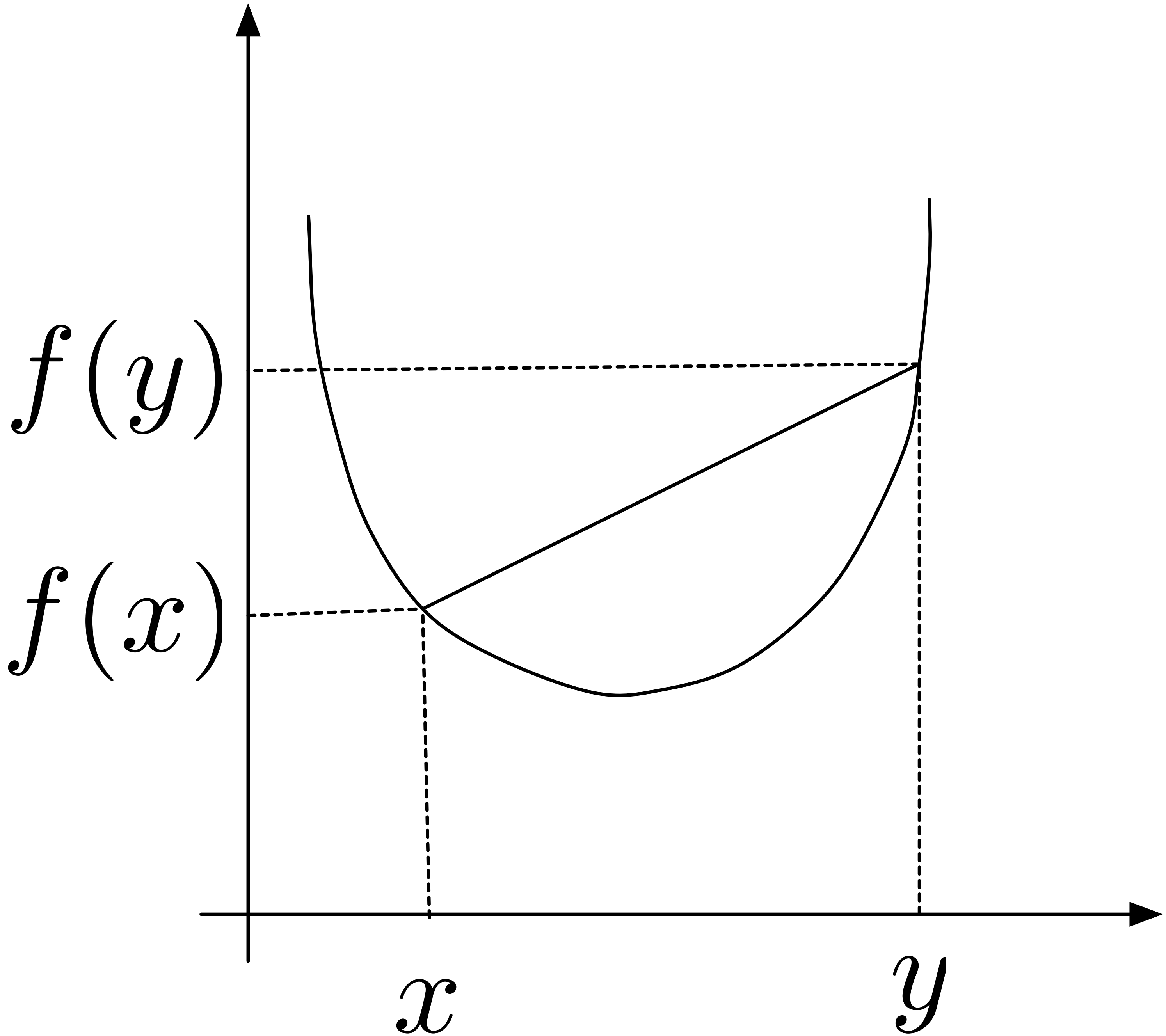
$$
\begin{aligned}
\text{minimize } & f_0(x) \\
\text{subject to } & f_i(x) \leq 0, \quad i = 1, 2, \ldots, m \\
& a_i^T x = b_i, \quad i = 1, 2, \ldots, p
\end{aligned}
$$

凸関数(convex function)

$f : \mathbb{R}^n \to \mathbb{R}$ が任意の $x, y \in \mathrm{dom} f$ について

$$f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$$

を満たすとき、f を凸関数(convex function)と呼ぶ。ただし、θ は $0 \le \theta \le 1$ を満たす実数。


$f(y)$
$f(x)$
x
y

凸関数：補足

- f がconvex(=凸) の場合、$-f$ はconcave という。

- 不等式

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y)$$

が成り立つとき、f は真に凸(strictly convex) である、という

凸関数の例 (1次元の場合)

- アフィン関数: $ax + b (a, b \in \mathbb{R})$
- 指数関数: $e^{ax} (a \in \mathbb{R})$
- ベキ関数: $x^{\alpha} (x \in \mathbb{R}_{++}, \alpha \geq 1$ or $\alpha \leq 0)$
- 負対数関数: $-\ln(x) (x \in \mathbb{R}_{++})$

クイズ

アフィン関数: $f(z) = az + b (a, b \in \mathbb{R})$ が凸関数であることを示せ。

解答例

$x, y \in \mathbb{R}(x \neq y)$ を任意に選ぶ。また、$\theta \in [0, 1]$ を任意に選ぶ。$z = \theta x + (1-\theta)y$ と置くと

$$\begin{aligned} f(z) &= a(\theta x + (1-\theta)y) + b && (1) \\ &= \theta a x + (1-\theta)a y + \theta b + (1-\theta)b && (2) \\ &= \theta(ax+b) + (1-\theta)(ay+b) && (3) \\ &= \theta f(x) + (1-\theta)f(y) && (4) \end{aligned}$$

が成り立つ。したがって、アフィン関数は凸関数である。

凸関数の例 (多次元の場合)

- アフィン関数:
  $ax^T + b(x \in \mathbb{R}^n, a \in \mathbb{R}^n, b \in \mathbb{R})$
- ノルム関数: $||x||(x \in \mathbb{R}^n)$

ノルム関数

以下では、$x \in \mathbb{R}^n$ を仮定する。

ℓ_1-ノルム

$$||x||_1 = |x_1| + |x_2| + \cdots |x_n|$$

ℓ_2-ノルム

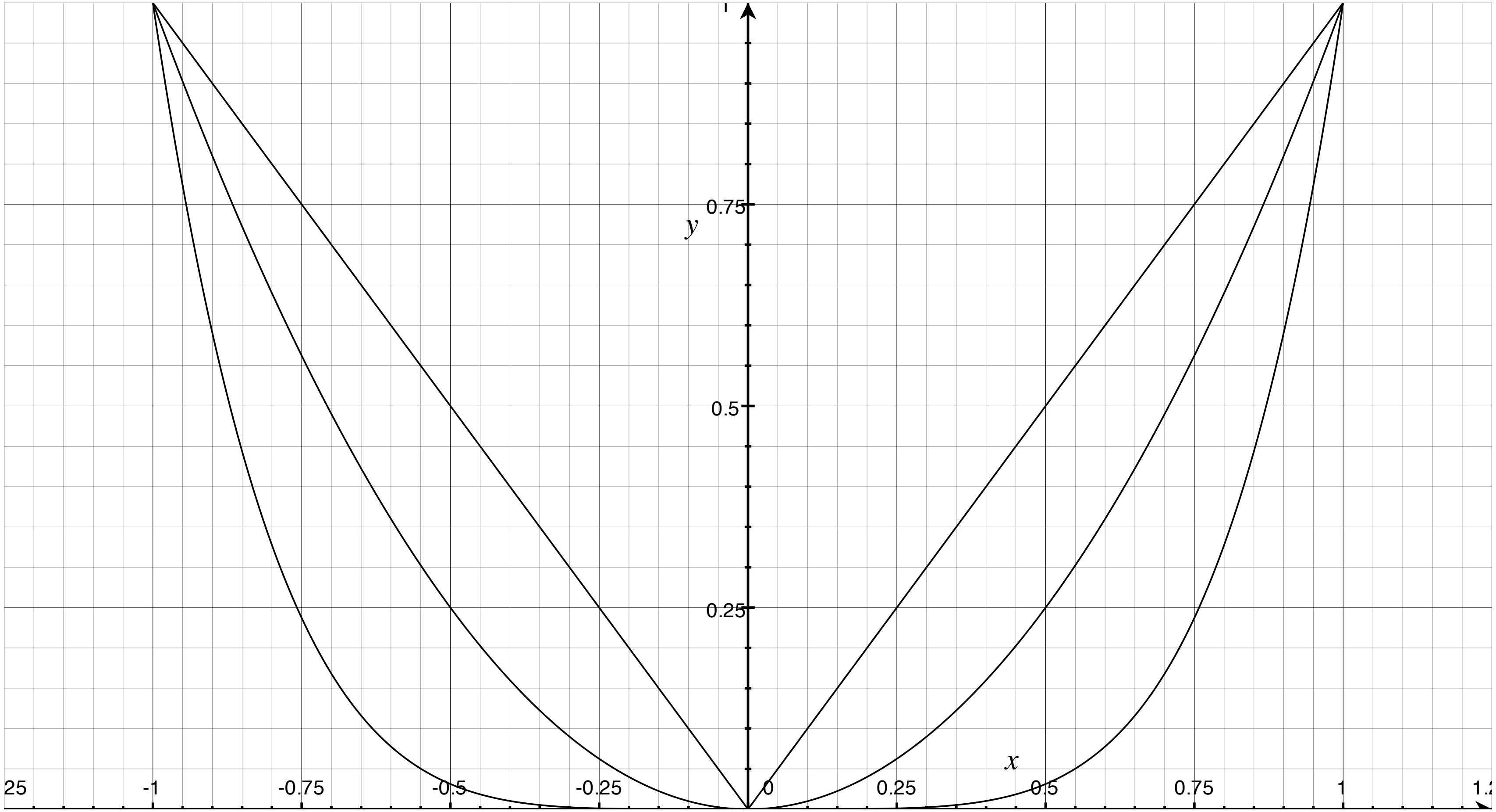
$$||x||_2 = \sqrt{|x_1|^2 + |x_2|^2 + \cdots |x_n|^2}$$

ℓ_∞-ノルム

$$||x||_\infty = \max\{|x_1|, |x_2|, \cdots |x_n|\}$$

ℓ_p-ノルム

$$||x||_p \equiv \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}, \quad p \geq 1$$



ノルム関数のプロット$(p = 1, 2, 5)$

クイズ

$x \in \mathbb{R}^2$ とする。ℓ_2-ノルム

$$f(x) = ||x||_2 = \sqrt{|x_1|^2 + |x_2|^2}$$

が凸関数であることを示せ。ただし、三角不等式

$$||a + b||_2 \leq ||a||_2 + ||b||_2 \tag{5}$$

を利用せよ。

解答例

$\alpha, \beta \in \mathbb{R}^2 (\alpha \neq \beta)$ を任意に選ぶ。また、
$\theta \in [0, 1]$ を任意に選ぶ。ここで、
$x = \theta\alpha + (1-\theta)\beta$ と置くと

$$f(x) = ||\theta\alpha + (1-\theta)\beta||_2 \tag{6}$$

となる。ここで三角不等式を利用すると

$$||\theta\alpha + (1-\theta)\beta||_2 \leq ||\theta\alpha||_2 + ||(1-\theta)\beta||_2$$

$$
\begin{aligned}
||\theta\alpha||_2 + ||(1-\theta)\beta||_2 &= \theta||\alpha||_2 + (1-\theta)||\beta||_2 \\
&= \theta f(\alpha) + (1-\theta) f(\beta)
\end{aligned}
$$

を得る。

凸関数の組み合わせ

- $\alpha, \beta \geq 0$としし、$f(x), g(x)$を凸関数$(x \in \mathbb{R}^n)$とすると

$$\alpha f(x) + \beta g(x)$$

  も凸関数となる(凸関数の非負係数和)。

- $f(x)$が凸関数の場合、$f(Ax + b)$も凸関数となる(アフィン関数との合成関数)。

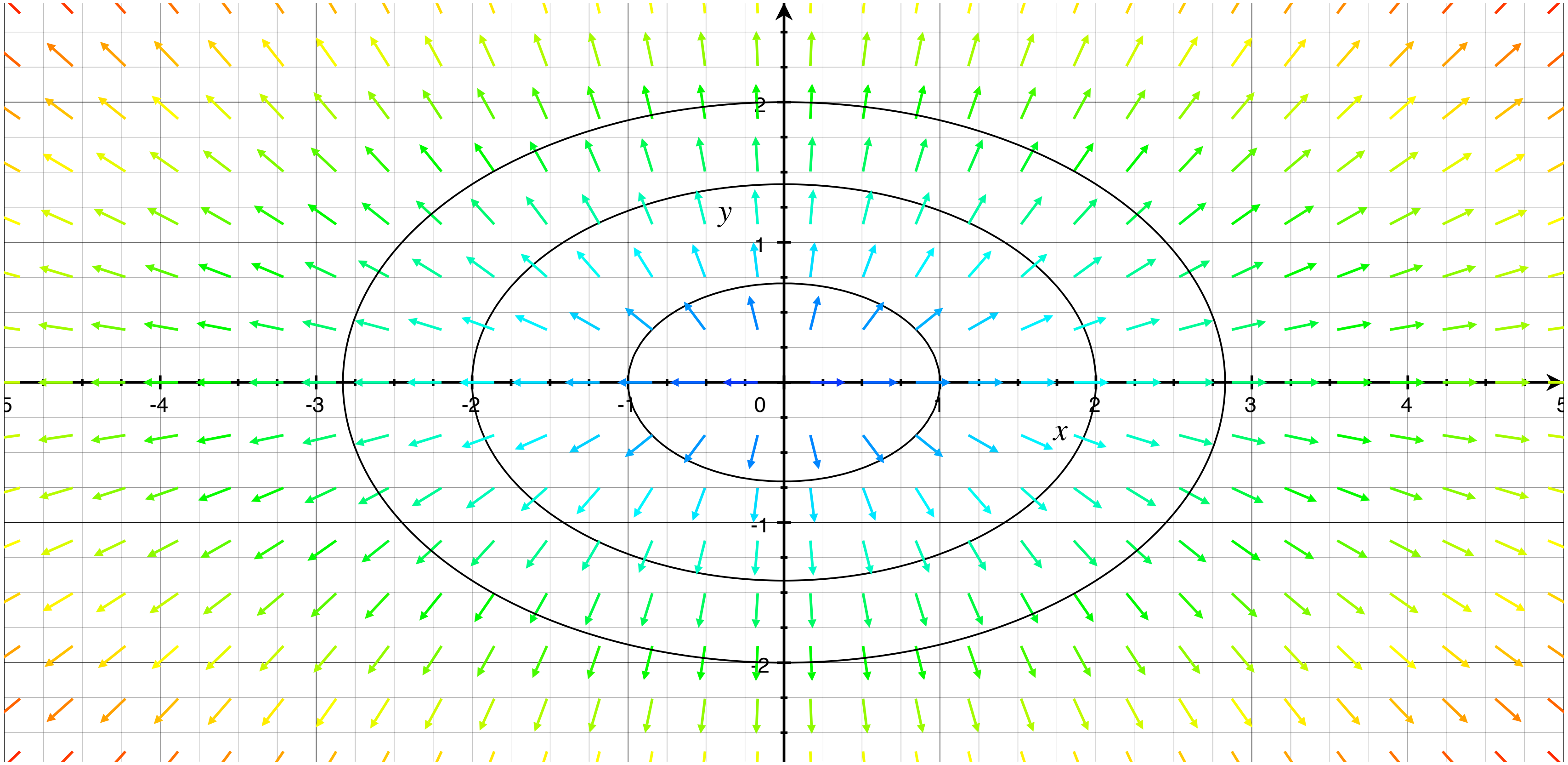
勾配ベクトル

f の勾配ベクトル(gradient):

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)^T$$

例：$f(x_1, x_2) = x^2 + 2y^2$ の場合、

$$\nabla f(x) = \begin{pmatrix} 2x \\ 4y \end{pmatrix}$$



関数 $f(x_1, x_2) = x^2 + 2y^2$ の勾配ベクトル場楕円は等高線 $x^2 + 2y^2 = q(q = 1, 4, 8)$ を表す。

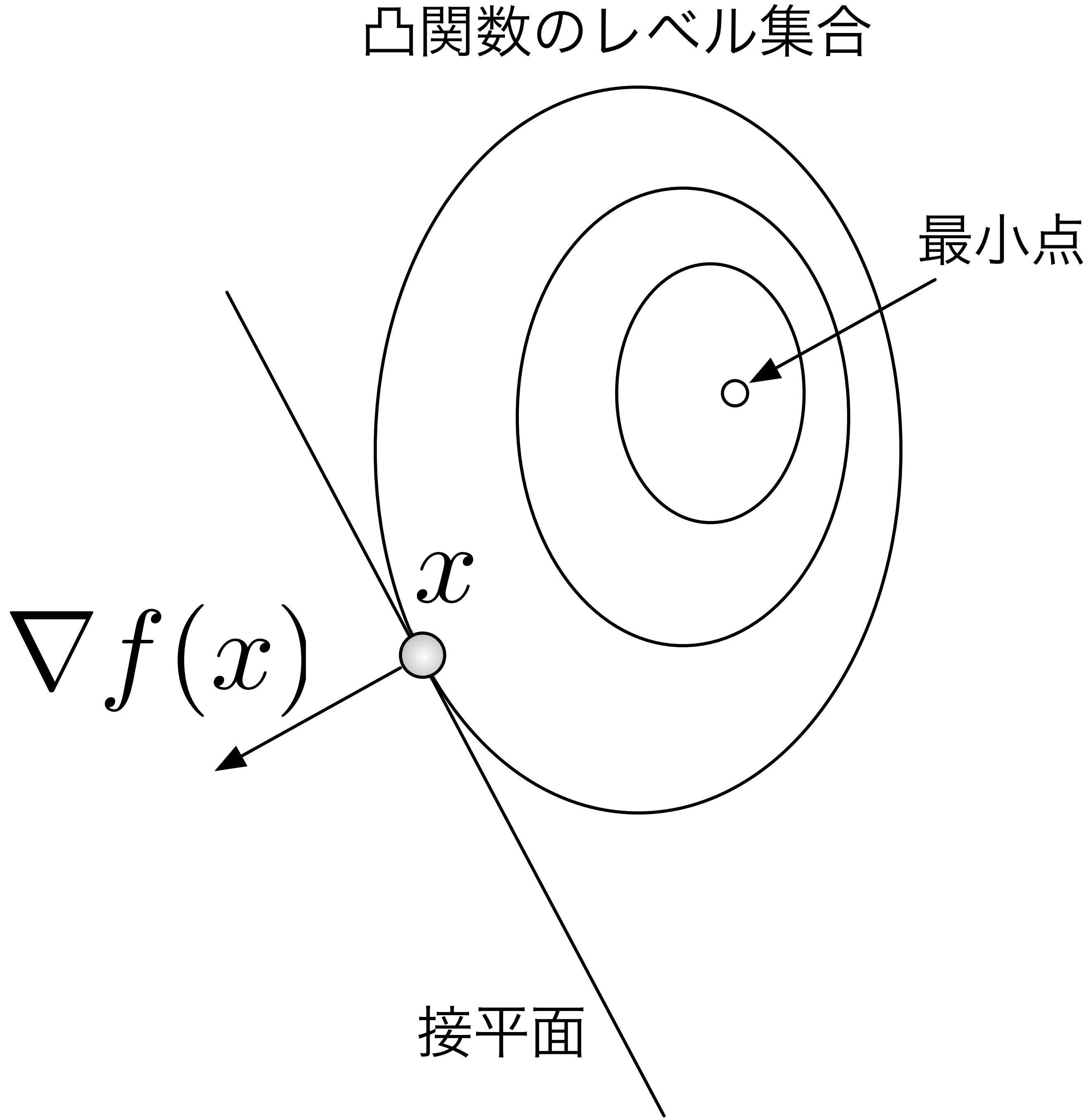
レベル集合と勾配ベクトル

$$S_\alpha = \{x \in \mathrm{dom} f : f(x) = \alpha\}$$

を f のレベル集合(＝等高線)と呼ぶ。

- 勾配ベクトルは、レベル集合の接平面の法線ベクトルとなっている。

例：天気図の場合、気圧の等高線がレベル集合となる。


凸関数のレベル集合
最小点
x
$\nabla f(x)$
接平面

2次の凸関数条件

$f : \mathbb{R}^n \to \mathbb{R}$, $\mathrm{dom}\, f$ が凸集合とする。

$$\nabla^2 f(x) \succeq 0, \quad x \in \mathrm{dom}\, f$$

と f が凸関数であることは同値である。
行列 A が $A \succeq 0$ であるとは、A が半正定値行列 (semi definite matrix) であることを意味する。

ヘッセ行列 (Hessian matrix)

$$\nabla^2 f(x) \equiv \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(x) & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} f(x) \\ \vdots & \vdots & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(x) & \cdots & \frac{\partial^2}{\partial x_n \partial x_n} f(x) \end{pmatrix}$$

半正定値行列(semidefinite matrix)

$n \times n$ 行列 $A \in \mathbb{R}^{n \times n}$ について

$$\forall x \in \mathbb{R}^n, \quad x^T A x \geq 0$$

が成立するとき、$A \succeq 0$ と書き、A を半正定値行列と呼ぶ。半正定値行列の固有値は、すべて非負である。