第 4 章　凸計画問題

凸計画は、線形計画問題や、Q が半正定値行列であるような 2 次計画問題などを含み、幅広い応用をもつ問題である。そればかりではなく、理論的にも良い性質をもち、さらに内点法を用いることにより、高速に解を得ることもできる。この章では、どのような問題が、凸計画問題となることかを紹介する。

4.1　凸計画問題

凸計画問題は理論的にも応用的にも重要な問題である。ここでは、まず、凸計画問題の重要性を議論する.

第 1 回目の講義でも述べたように凸計画問題とは、目的関数が凸関数で、制約条件が凸集合である数理計画問題である。

$$
\begin{array}{ll}
\min & f(x) \\
\text{subjectoto} & x \in \mathcal{F}
\end{array}
$$

ここで、f が凸関数であり、$\mathcal{F}$ が凸集合である。凸計画問題に対しては、次のことが知られている。

- 線形計画問題や凸 2 次計画問題など、多くの問題が凸計画問題として定式化できる。
- 凸計画問題において、最適性の条件を満たす点は、その問題の大域的最適解となる。
- 凸計画問題のラグランジュ双対問題は、元の凸計画問題と等価になる (同じ最適値を持つ)。
- 内点法を用いることにより、高速に解を得ることができる。

1 番目の項目は、凸計画問題が幅広い応用を持つことを意味している。2 番目の項目は、特に重要である。数理計画問題の多くのアルゴリズムでは、その大域的最適解ではなく、最適性の条件を満たす点しか求めることができない。そのため、そのような点が大域的最適解であることが保証できている問題であるということは、実用上に非常に重要なことである。3 番目の項目も、数理計画問題を考察する上で重要である。双対問題は、元の問題に比べて解きやすい場合がある。また、反復法によって、元の問題と双対問題を同時に解けば、それぞれの目的関数の値の差を見ることによって、その反復点がどれくらいよい解であるかを判断することができる。なお、ラグランジュ双対問題については、次回紹介する。問題を解く観点から言えば、内点法を用いることによって、一般の数理計画問題に比べて、凸計画問題は効率的に解くことができる。

それでは、凸計画問題になるのは、どのような問題 $(f, g, h, \ldots)$ であろうか？そのためには、どのようなときに実行可能集合は凸集合になるのか、また、どのようなときに目的関数は凸関数になるのかを知る必要がある。次節以降では、凸に関連したいくつかの性質を紹介する。なお、関数 f が凸関数であるとは,

$$
f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y) \quad \forall x, y \in R^n,\ \forall \alpha \in [0,1] \tag{4.1}
$$

が成り立つことをいい，集合 S が凸集合であるとは，

$$\forall x, y \in S,\ \ \forall \alpha \in [0,1] \Rightarrow \alpha x + (1-\alpha)y \in S \tag{4.2}$$

が成り立つことであったことを思い出そう。

4.2　凸集合となる条件

ここでは、どのようなときに実行可能集合 $\mathcal{F}$ が凸集合となるのか調べる。

まず、不等式制約が凸集合となる条件を与える。

(a)

凸集合の定義 (4.2) より、任意の $x, y \in S$ と $\alpha \in [0,1]$ に対して、$\alpha x + (1-\alpha)y \in S$ となることを示せばよい。$x, y \in S$ であることから、$g(x) \le 0$ かつ $g(y) \le 0$ である。さらに、g が凸関数であることと $\alpha \in [0,1]$ であることより、

$$g(\alpha x + (1-\alpha)y) \le \alpha g(x) + (1-\alpha)g(y) \le 0$$

となる。この式は $\alpha x + (1-\alpha)y \in S$ であることを意味している。

なお、不等式制約の実行可能集合 S が凸集合であっても、g が凸関数となるとは限らない。例えば、x^3 は凸関数ではない。しかしながら、$x^3 \le 0$ は $\{x \in R \mid x \le 0\}$ となり凸集合となる。

次に等式制約が凸集合となる条件を考える。

(b)

このことも容易に示すことができる。h が $R^n \to R$ のアフィン関数であるので、ある n 次元ベクトル a とスカラー b を用いて、$h(x) = a^T x + b$ と表すことができる。$x, y \in S$ と $\alpha \in R$ に対して、

$$h(\alpha x + (1-\alpha)y) = a^T(\alpha x + (1-\alpha)y) + (\alpha + (1-\alpha))b(\alpha h(x) + (1-\alpha)h(y) = 0$$

となる。これは $\alpha x + (1-\alpha)y \in S$、つまり S が凸集合であることをあらわしている。なお、この関係の逆は必ずしもなりたたない。実際、$h(x) = \max\{0, x\}$ とすると、$S = \{x \in R \mid x \le 0\}$ となり、S は凸集合となる。しかしながら、h は線形関数ではない。

一般の数理計画問題の実行可能集合は、複数の不等式制約や等式制約を満たす点の集合としてあらわされる。そこで、凸集合の共通集合が、どのような集合となるかを知ることは重要である。

(c)

ここで、$x, y \in S \cap T$ として、$\alpha \in [0,1]$ とする。$x, y \in S$ であることから、$\alpha x + (1-\alpha)y \in S$ である。同様にして、$\alpha x + (1-\alpha)y \in T$ となる。つまり、$\alpha x + (1-\alpha)y \in S \cap T$ となるので、$S \cap T$ は凸集合である (図 4.1)。

以上の (a), (b), (c) をまとめると、実行可能集合 $\mathcal{F}$ が凸集合となる十分条件が以下のように表される。

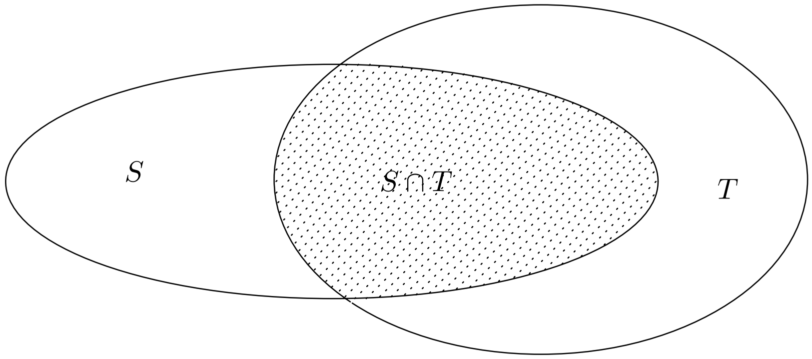

図 4.1: 凸集合の共通集合

$h : R^n \to R^m$ はアフィン関数とし、$g_i, i = 1, \ldots, r$ は凸関数とする。さらに $X \subseteq R^n$ を凸集合とする。このとき、

$$\mathcal{F} := \{x \in X \mid h(x) = 0, g(x) \leq 0\}$$

は凸集合である。

なお、X が離散的に表現されている組み合せ最適化問題は凸計画問題にならない。

4.3　凸関数となる条件

前節では、不等式制約にもちいる関数 g_i が凸関数となることが実行可能集合が凸集合となるための十分条件であることをみた。また、凸計画問題であるためには、目的関数 f が凸関数であるかどうか知る必要がある。この節では、凸関数となる条件をいくつか紹介する。

まず、一回微分可能な関数が凸関数となるための必要十分条件は以下のように与えられる。

(i)

この性質は図 4.2 を見れば明らかである。これは凸関数の接線 (2 次元以上であれば超平面) が、その関数よりも下にあることを意味している。実際に証明することも簡単である。f が凸関数であることより、任意の $\alpha \in (0, 1]$ に対して、

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$$

が成り立つ．この式を変形すると，

$$\frac{f(y + \alpha(x-y)) - f(y)}{\alpha} \leq f(\bar{x}) - f(y)$$

を得る．この式において，$\alpha \to 0$ を考えると、

$$\lim_{\alpha \to 0} \frac{f(y + \alpha(x-y)) - f(y)}{\alpha} \leq f(\bar{x}) - f(y) \tag{4.3}$$

となる。ここで、$p : R \to R^n$ を $p(\alpha) = y + \alpha(x-y)$ とし、$q : R \to R$ を p と f の合成関数、つまり $q(\alpha) = f(h(\alpha))$ とする。このとき、(4.3) の左辺は、

$$\lim_{\alpha \to 0} \frac{f(y + \alpha(x-y)) - f(y)}{\alpha} = \lim_{\alpha \to 0} \frac{q(\alpha)) - q(0)}{\alpha} = \nabla q(0)$$

図 4.2: 凸関数と接線

となる。合成関数の微分より、

$$\nabla q(0) = \nabla h(0)^T \nabla f(h(0)) = (x-y)^T \nabla f(y) = \nabla f(y)^T (x-y)$$

を得る。この式と (4.3) より、求めたい不等式が得られる。逆に、性質 (i) の不等式が成り立つとする。ここで $\bar{x}, \bar{y} \in R^n$ と $\alpha \in [0,1]$ とする。まず、性質 (i) の不等式に $x = \bar{x}$ と $y = (1-\alpha)\bar{x} + \alpha\bar{y}$ を代入すると、

$$f(\bar{x}) - f((1-\alpha)\bar{x} + \alpha\bar{y}) \geq \nabla f((1-\alpha)\bar{x} + \alpha\bar{y})^T (\bar{x} - ((1-\alpha)\bar{x} + \alpha\bar{y})) \tag{4.4}$$

を得る。同様に $x =$ y と $y = (1-\alpha)\bar{x} + \alpha\bar{y}$ を代入すると、

$$f(\bar{y}) - f((1-\alpha)\bar{x} + \alpha\bar{y}) \geq \nabla f((1-\alpha)\bar{x} + \alpha\bar{y})^T (\bar{y} - ((1-\alpha)\bar{x} + \alpha\bar{y})) \tag{4.5}$$

を得る。ここで、(4.4) の両辺を $1-\alpha$ 倍した不等式と、(4.5) の両辺を α 倍した不等式を足し合わせると、

$$\begin{aligned} & (1-\alpha) f(\bar{x}) + \alpha f(\bar{y}) - f((1-\alpha)\bar{x} + \alpha\bar{y}) \\ \geq \;& \nabla f((1-\alpha)\bar{x} + \alpha\bar{y})^T ((1-\alpha)\bar{x} + \alpha\bar{y} - ((1-\alpha)\bar{x} + \alpha\bar{y})) = 0 \end{aligned}$$

となる。この式は、f が凸関数であることを示している。

次に関数 f が 2 回微分可能な場合をみてみよう。

(ii)

この性質を証明することは多少準備が必要となるので、省略する。この性質では、2 次関数 $ax^2 +$

$bx+c$ が凸関数となるための必要十分条件は $a \geq 0$ であることを意味している。次にこのことを n 次元に拡張した場合、つまり $f: R^n \to R$ が、$n \times n$ の対称行列 Q と n 次元ベクトル q によって、$f(x) = \frac{1}{2}x^TQx + q^Tx$ と表されている場合を考えてみよう。この関数のヘッセ行列は $\nabla^2 f(x) = Q$ である。このとき、Q が半正定値行列であれば、f が凸関数となることを示してみよう。これは、$x, y \in R^n$ と $\alpha \in [0,1]$ に対して、

$$\alpha x^TQx + (1-\alpha)y^TQy - (\alpha x + (1-\alpha)y)^TQ(\alpha x + (1-\alpha)y) \geq 0$$

となることを示せばよい。この式の左辺は

$$\begin{aligned}
&\alpha x^TQx + (1-\alpha)y^TQy - (\alpha x + (1-\alpha)y)^TQ(\alpha x + (1-\alpha)y) \\
&= \alpha(1-\alpha)(x^TQx + y^TQy - 2x^TQy) \\
&= \alpha(1-\alpha)(x-y)^TQ(x-y) \\
&\geq 0
\end{aligned}$$

となる。ここで 1 番目の等式は Q が対称行列であること、最後の不等式には Q が半正定値行列であり、$\alpha \in [0,1]$ であることを用いている。

この性質 (ii) は与えられて関数が凸関数であるかをチェックする上でよく用いられる。例えば、$f(x) = e^{x_1+x_2}$ という関数が凸関数かどうか調べるとき、この勾配は

$$\nabla f(x) = \begin{pmatrix} e^{x_1+x_2} \\ e^{x_1+x_2} \end{pmatrix}$$

となる。ヘッセ行列は、

$$\nabla^2 f(x) = \begin{pmatrix} e^{x_1+x_2} & e^{x_1+x_2} \\ e^{x_1+x_2} & e^{x_1+x_2} \end{pmatrix}$$

となる。この行列は任意の $v \in R^2$ に対して、

$$(v_1 \;\; v_2)\begin{pmatrix} e^{x_1+x_2} & e^{x_1+x_2} \\ e^{x_1+x_2} & e^{x_1+x_2} \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = e^{x_1+x_2}(v_1+v_2)^2$$

となり、任意の $x \in R^2$ に対して非負である。よって、ヘッセ行列は半正定値行列となるので、この f は凸関数である。また、$Q = 0$ は半正定値行列であるため、線形関数 $f(x) = c^Tx$ も凸関数となる。このことより、線形計画問題や 2 次計画問題の実行可能集合

$$calF := \{x \in R^n \mid Ax = b, x \geq 0\}$$

が凸集合となることがわかる。さらに、線形計画問題の目的関数 c^Tx は凸関数であるので、線形計画問題は凸計画問題である。また、Q が対称行列であるとき、$f(x) = \frac{1}{2}x^TQx + q^Tx$ のヘッセ行列は Q となるので、Q が半正定値行列であれば、f は凸関数となる。このため、Q が半正定値行列である 2 次計画問題は凸計画問題となる。

微分可能な関数に対しては、上記の (i), (ii) を用いることによって、凸関数であるかどうか判断することができた。そこで、次に、関数の和や合成が凸関数となるための条件を調べることにしよう。

(iii)

このことも定義に当てはめれば容易に確かめることができる。実際、f_1 と f_2 を凸関数とすると、

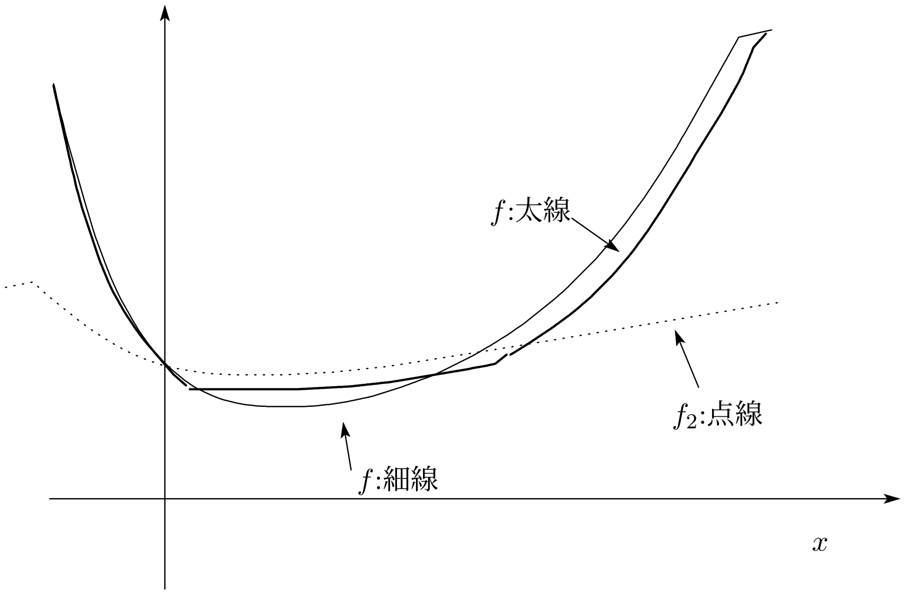

図 4.3: max 関数の凸性

その和である $f(x) := f_1(x) + f_2(x)$ は、

$$\begin{aligned}\alpha f(x)+(1-\alpha)f(y) &= \alpha f_1(x)+(1-\alpha)f_1(y)+\alpha f_2(x)+(1-\alpha)f_2(y)\\ &\geq f_1(\alpha x+(1-\alpha)y)+f_2(\alpha x+(1-\alpha)y)\\ &= f(\alpha x+(1-\alpha)y)\end{aligned}$$

となるので、凸関数である。このため、$f(x_1, x_2) = e^{x_1+x_2} + x_1 + x_2$ は凸関数となる。

(iv)

ここで単調増加とは、$b \geq 0$ であるとき、a から b 増加すると関数値も増加、つまり $f(a) \leq f(a+b)$ となることである。この性質 (iv) も定義にあてはめれば、

$$h(f(\alpha x + (1-\alpha y)) \leq h(\alpha f(x) + (1-\alpha)f(y)) \leq \alpha h(f(x)) + (1-\alpha)h(f(x))$$

となることよりわかる。ここで、1 番目の不等式には h が単調増加関数であることと f が凸関数であることを用いた。また最後の不等式は h が凸関数であることを用いている。

(v)

この性質も定義に当てはめれば、容易に導くことができる。ここで、任意の $x, y \in R^n$ と $\alpha \in [0,1]$ をとる。まず、$f_1(\alpha x+(1-\alpha)y) \leq f_2(\alpha x+(1-\alpha)y)$ の場合を考える。このとき、f_2 の凸性より、

$$\begin{aligned}f(\alpha x+(1-\alpha)y) &= f_2(\alpha x+(1-\alpha)y)\\ &\leq \alpha f_2(x)+(1-\alpha)f_2(y)\\ &\leq \alpha\max\{f_1(x),f_2(x)\}+(1-\alpha)\max\{f_1(y),f_2(y)\}\\ &= \alpha f(x)+(1-\alpha)f(y)\end{aligned}$$

となる。同様にして、$f_1(\alpha x + (1-\alpha)y) > f_2(\alpha x + (1-\alpha)y)$ の場合も示すことができる。

練習問題 1 $\max\{0, e^{x_1+x_2} - x_1\}^2$ が凸関数となることを示せ。

最後に凸関数となるいくつかの良く知られた関数を紹介しよう。これらの関数と性質 (iii)(iv) を組み合せることによって、多くの凸関数を生み出したり、また凸関数であることをチェックすることができる。

以下の関数は凸関数である。

- 線形関数
- max 関数
- $-\log$, 指数関数。これらの関数は単調増加な凸関数でもある。
- 半正定値行列を用いた 2 次関数