最適化理論

ラグランジアンと双対関数

ラグランジアンと双対関数

一般の数理計画問題

かならずしも凸計画とは限らない、次の数理計画問題を考える。

$$
\begin{aligned}
\text{minimize}\ & f_0(x) \\
\text{subject to}\ & f_i(x) \le 0, \quad i = 1, 2, \ldots, m \\
& h_i(x) = 0, \quad i = 1, 2, \ldots, p
\end{aligned}
$$

ただし、$x \in \mathbb{R}^n$ である。この問題の最適値を p^* とし、実行可能領域を F とする。

ラグランジアンの定義

$x \in \mathbb{R}^n, \lambda \in \mathbb{R}^m, \nu \in \mathbb{R}^p$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

- λ_i: $f_i(x) \le 0$ に対応するラグランジュ乗数
- ν_i: $h_i(x) = 0$ に対応するラグランジュ乗数

ラグランジュ双対関数

$$
\begin{aligned}
g(\lambda, \nu) &= \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \\
&= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)
\end{aligned}
$$

ここで、

$$
\mathcal{D} = \bigcap_{i=0}^{m} \mathrm{dom} f_i \cap \bigcap_{i=1}^{p} \mathrm{dom} h_i
$$

ラグランジュ双対関数の性質

(1)Concave 関数となる。

(2)$\lambda \geq 0$のとき、$g(\lambda, \nu) \leq p^*$ が成り立つ(最適値の下界値)。

(証明) 任意の$\tilde{x} \in F$を仮定する。

$$f_0(\tilde{x}) \geq L(\tilde{x}, \lambda, \nu) \geq g(\lambda, \nu)$$

最適点もF内に含まれるので、上の不等式が適用できる。

具体例

次の凸計画問題を考える。ここで、$x \in \mathbb{R}$ である。

$$\text{minimize } x^2$$
$$\text{subject to } -x+1 \leq 0$$

- この問題のラグランジアンを書き下せ。
- ラグランジュ双対関数を求めよ。
- ラグランジュ双対関数が最適値の下界となっていることを確認せよ。

ラグランジアンは

$$L(x, \lambda, \nu) = x^2 + \lambda(-x + 1)$$

となる。L を x の関数としてみると2次関数であり、凸である。したがって、極値条件

$$\frac{\partial L}{\partial x} = 2x - \lambda = 0$$

より、$x = \lambda/2$ を得る。この値を L に代入することにより、ラグランジュ双対関数が

$$
\begin{aligned}
g(\lambda, \nu) &= \left(\frac{\lambda}{2}\right)^2 + \lambda\left(-\frac{\lambda}{2} + 1\right) \\
&= -\frac{\lambda^2}{4} + \lambda = -\frac{1}{4}(\lambda - 2)^2 + 1
\end{aligned}
$$

と得られる。$g(\lambda, \nu)$ は $\lambda = 2$ のとき、極大値(最大値)を取り最大値 1 である。確かに

$$
p^* = 1 \geq g(2, \nu) = 1
$$

ラグランジュ双対問題

下界$p^* \geq g(\lambda, \nu)$ を最も厳しくするには、$g(\lambda, \nu)$ を最大化すればよい。

ラグランジュ双対問題

$$\begin{aligned} &\text{maximize } g(\lambda, \nu) \\ &\text{subject to } \lambda \geq 0 \end{aligned}$$

- もとの最適化問題を主問題(primal problem)と呼ぶ。
- x を主変数、(λ, ν) を双対変数と呼ぶ。

弱双対性と強双対性

双対問題の最適値をd^*とする。

<u>弱双対性</u>　$d^* \le p^*$が主問題が凸であっても非凸であっても成り立つ。

<u>強双対性</u>　$d^* = p^*$が多くの凸計画問題(すべてではない）について成り立つ。

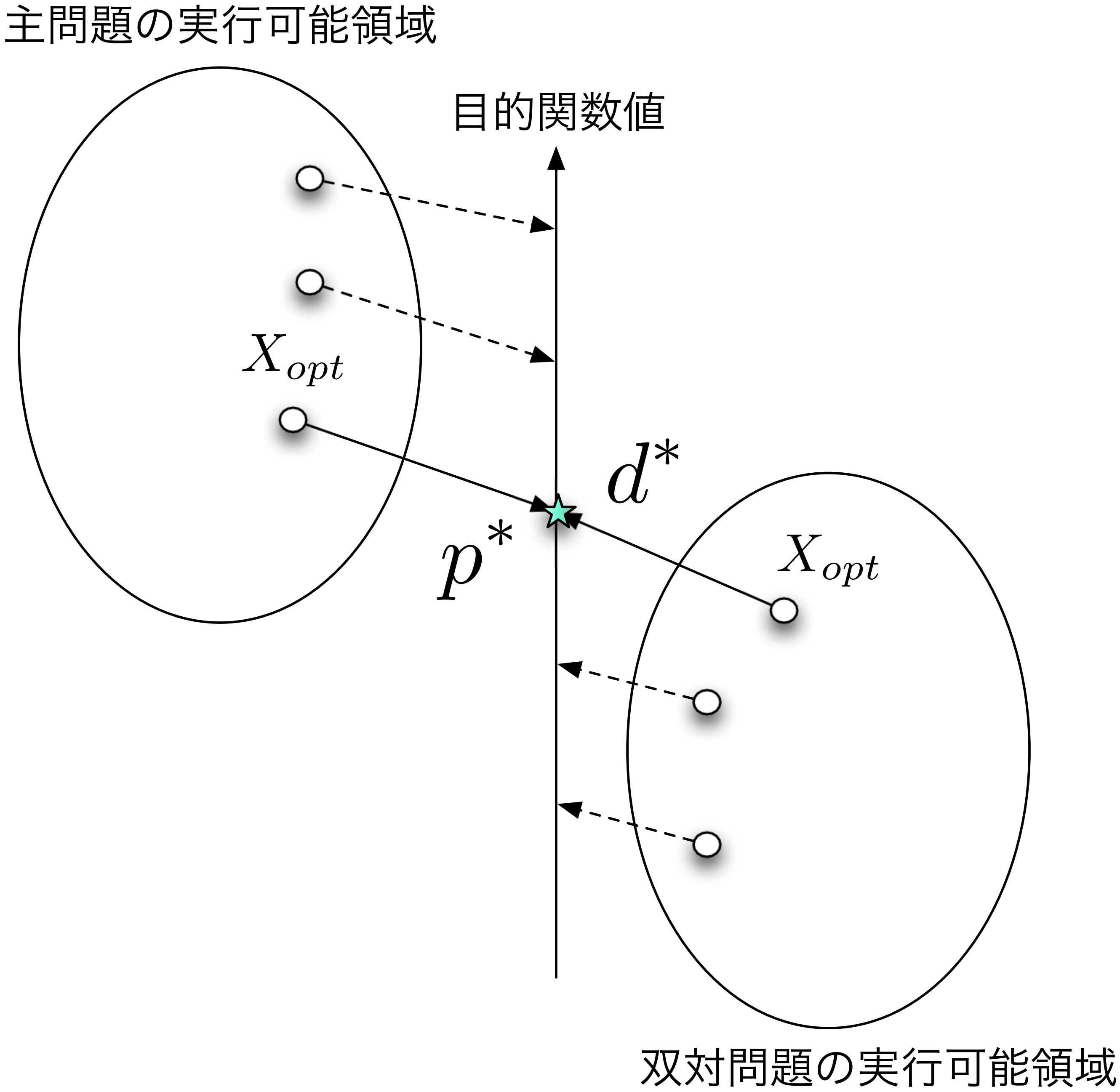
強双対性が成り立っている例

主問題 ($p^* = 1$)

$$\text{minimize } x^2$$
$$\text{subject to } -x+1 \leq 0$$

双対問題($d^* = 1$)

$$\text{maximize } -\frac{1}{4}\lambda^2 + \lambda$$
$$\text{subject to } \lambda \geq 0$$


主問題の実行可能領域
目的関数値
X_{opt}
d^*
p^*
X_{opt}
双対問題の実行可能領域

KKT条件

- KKT条件(Karush-Kuhn-Tucker)条件は、強双対性の成立する凸計画問題の最適解が満たすべき必要十分条件である。
- KKT条件は、最適解の解析的な導出や数値最適化アルゴリズムの原理としても重要である。

相補性条件

x^* を主問題の最適解、λ^* と ν^* を双対問題の最適解とする。強双対性を仮定すると

$$
\begin{aligned}
f_0(x^*) = g(\lambda^*, \nu^*) &= \inf_{x \in \mathcal{D}} L(x, \lambda^*, \nu^*) \\
&\leq f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{i=1}^{m} \nu_i^* f_i(x^*) \\
&\leq f_0(x^*)
\end{aligned}
$$

すなわち、上の不等式部分は全て等号となる。

この等式より次のことが言える。

- $\sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$となる。

  $$\lambda_i > 0 \Rightarrow f_i(x^*) = 0, \quad \lambda_i = 0 \Rightarrow f_i(x^*) < 0$$

  この条件を相補性条件という。

- $f_0(x^*) = \inf_{x \in \mathcal{D}} L(x, \lambda^*, \nu^*)$となることから、

  $$\nabla_x L(x, \lambda^*, \nu^*) \mid_{x=x^*} = 0$$

KKT条件

条件1　xが主問題の実行可能解であること。

$$f_i(x) \leq 0, i = 1, ..., m, \quad h_i(x) = 0, i = 1, ..., p$$

条件2　双対変数の制約: $\lambda \geq 0$

条件3　相補性条件: $\lambda_i f_i(x) = 0 \ (i = 1, \ldots, m)$

条件4　ラグランジアンのxに関する極値条件:

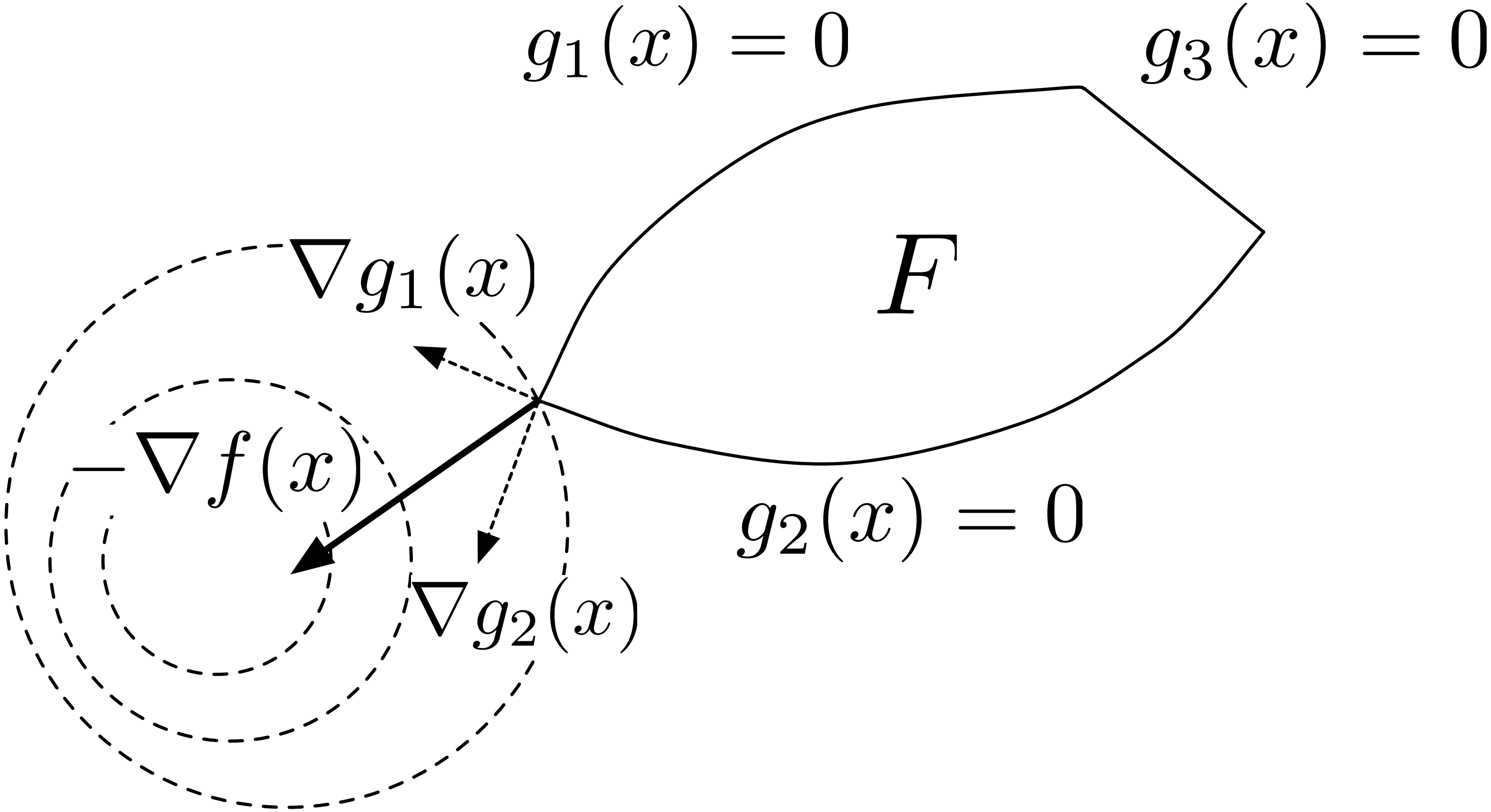
$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0$$

KKT条件の役割

強双対性が成り立つ凸計画問題の場合、
(x, λ, ν) が最適値 $\Leftrightarrow$ KKT条件が成り立つ

- 非凸な問題でも強双対性が成り立つ場合、“(x, λ, ν) が最適値 $\Rightarrow$ KKT条件成立” となるが逆は必ずしも真ならず。
- 凸計画問題の場合、KKT条件を解析的、数値的に解くことによっても最適解を求めることができる。

KKT条件の図的解釈



$$-\nabla f(x) = \lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x) + 0 \cdot \nabla g_3(x)$$

練習問題

次の問題をKKT条件を利用して最適解を求めよ。

$$
\begin{aligned}
&\text{minimize } x^2 \\
&\text{subject to } -x+1 \le 0
\end{aligned}
$$

すでに見たようにこの問題のラグランジアンは

$$L(x, \lambda, \nu) = x^2 + \lambda(-x + 1)$$

となる。このとき、条件4は、

$$\nabla L(x, \lambda, \nu) = 2x - \lambda = 0$$

となる。また、条件3(相補性条件)より、

$$\lambda(-x + 1) = 0$$

これを解くと $\lambda = 0$ または、$x = 1$ となるが、$\lambda = 0, x = 0$ となるため、x の実行可能条件を満たさない。

一方、$x = 1, \lambda = 2$ はKKT条件のすべての条件を満足するので、これが解である。