離散最適化基礎論

第 13 回

ラグランジュ緩和 (2)：最適ラグランジュ緩和

岡本 吉央　(電気通信大学)

okamotoy@uec.ac.jp

2023 年 1 月 17 日

最終更新：2023 年 1 月 23 日　00:46

講義計画 (前半)

＜準備＞

1. 整数計画法と線形計画法 (10/4)
2. 線形計画法の復習 (1)：線形不等式系と凸多面集合 (10/11)

∗ 休み (体育祭) (10/18)

3. 線形計画法の復習 (2)：単体法と双対定理 (10/25)
4. 線形計画緩和 (11/1)

＜モデリング＞

5. 整数計画モデリング (1)：組合せ最適化問題 (11/8)
6. 整数計画モデリング (2)：より複雑な問題 (11/15)
7. 整数計画モデリング (3)：離接計画 (11/22)

講義計画 (後半)

＜アルゴリズム＞

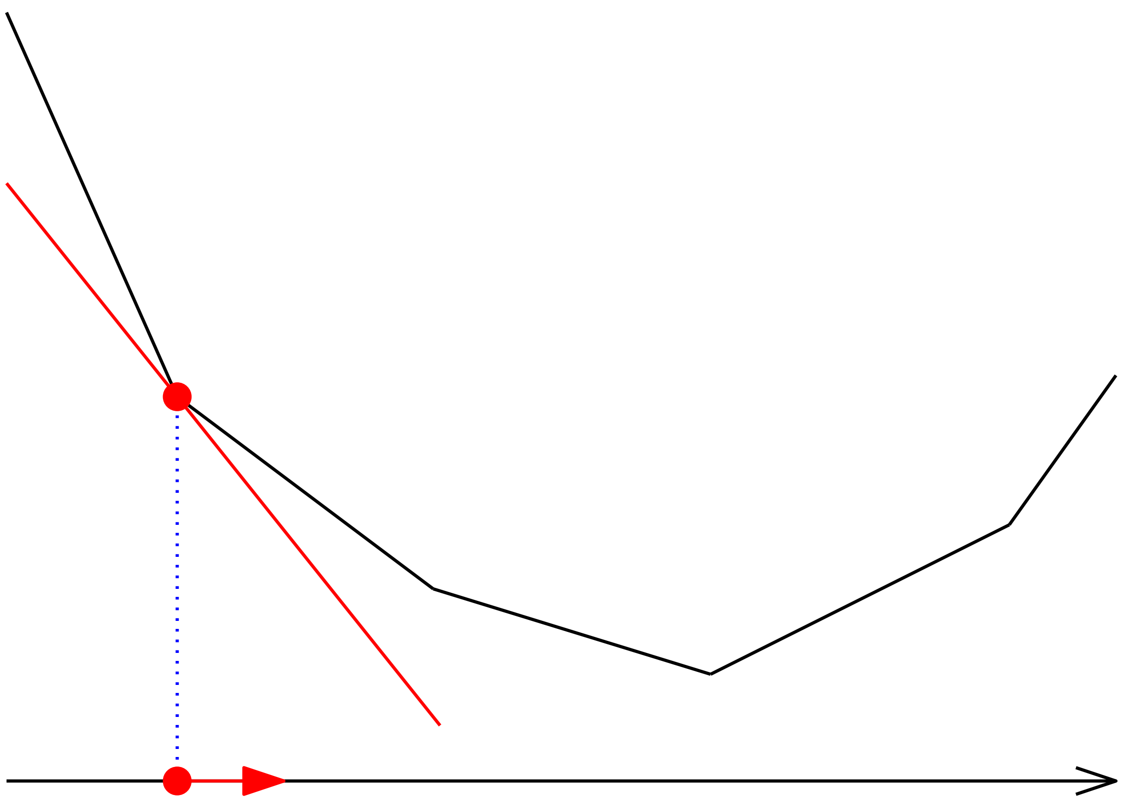
8.	分枝限定法	(11/29)
9.	切除平面法	(12/6)
10.	妥当不等式の追加	(12/13)
11.	列生成法	(12/20)
*	休み (国内出張)	(12/27)
*	休み (冬季休業)	(1/3)
12.	ラグランジュ緩和 (1)：原理	(1/10)
13.	ラグランジュ緩和 (2)：最適ラグランジュ緩和	(1/17)
14.	前求解	(1/24)

＜まとめ＞

15.	期末試験	(1/31)

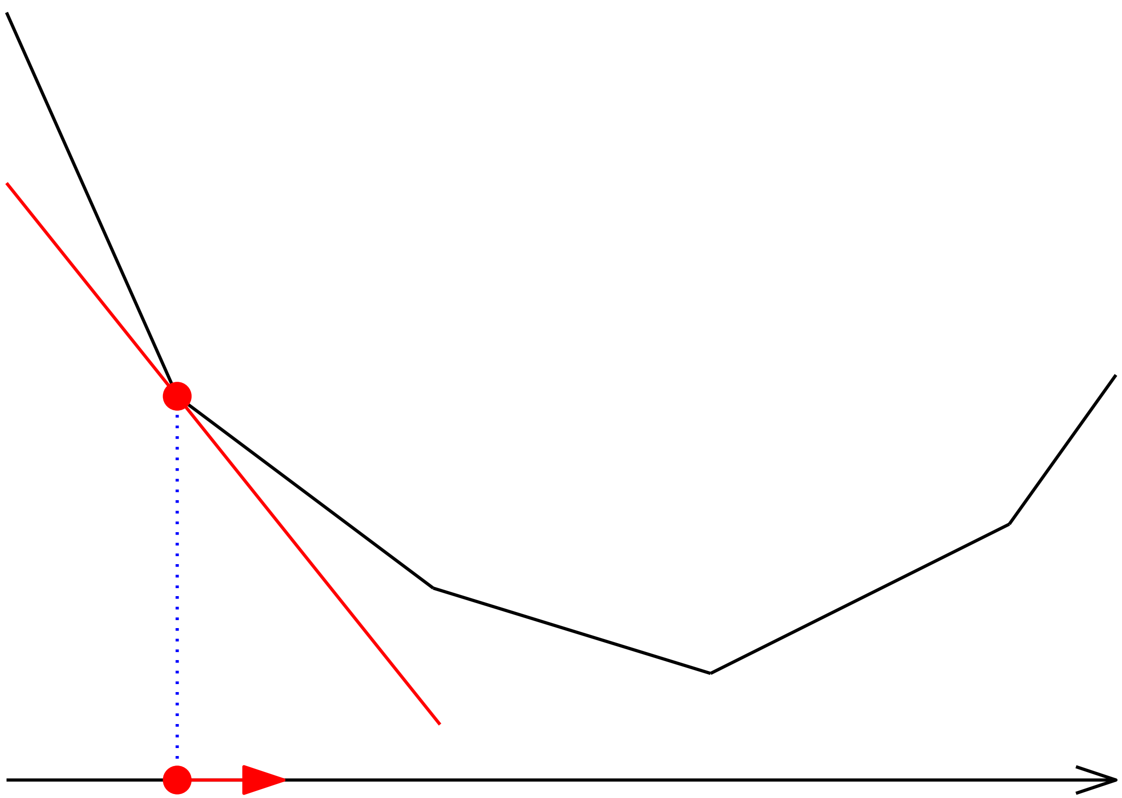
今日の内容

- ラグランジュ緩和：復習
- ラグランジュ緩和の性質
- 劣微分と劣勾配
- 劣勾配法



今日の内容

- ラグランジュ緩和：復習
- ラグランジュ緩和の性質
- 劣微分と劣勾配
- 劣勾配法



整数計画問題 P

$$
\begin{aligned}
\text{min.}\quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\
\text{s.t.}\quad & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\
& A_2\boldsymbol{x} = \boldsymbol{b}_2, \\
& \boldsymbol{x} \geq \boldsymbol{0}, \\
& \boldsymbol{x} \in \mathbb{Z}^n
\end{aligned}
$$

変数： $\boldsymbol{x} \in \mathbb{R}^n$

定数： $A_1 \in \mathbb{R}^{m_1 \times n}, A_2 \in \mathbb{R}^{m_2 \times n}, \boldsymbol{b}_1 \in \mathbb{R}^{m_1}, \boldsymbol{b}_2 \in \mathbb{R}^{m_2}, \boldsymbol{c} \in \mathbb{R}^n$

ラグランジュ緩和：定義 (等式標準形)

整数計画問題 P $\longrightarrow$ ラグランジュ緩和 L($\boldsymbol{\lambda}$)

$$\begin{aligned} \text{min.}\quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{s.t.}\quad & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & A_2\boldsymbol{x} = \boldsymbol{b}_2, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{aligned}$$

$$\begin{aligned} \text{min.}\quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \\ \text{s.t.}\quad & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{aligned}$$

(Lagrangian relaxation)

変数： $\boldsymbol{x} \in \mathbb{R}^n$

定数： $A_1 \in \mathbb{R}^{m_1 \times n}, A_2 \in \mathbb{R}^{m_2 \times n}, \boldsymbol{b}_1 \in \mathbb{R}^{m_1}, \boldsymbol{b}_2 \in \mathbb{R}^{m_2}, \boldsymbol{c} \in \mathbb{R}^n$

$\boldsymbol{\lambda} \in \mathbb{R}^{m_2}$

直感

P における制約の非充足 $\to$ L($\boldsymbol{\lambda}$) における罰則

ラグランジュ緩和：性質

整数計画問題 P $\longrightarrow$ ラグランジュ緩和 L($\boldsymbol{\lambda}$)

$$\begin{aligned} \text{min.} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{s.t.} \quad & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & A_2\boldsymbol{x} = \boldsymbol{b}_2, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{aligned}$$

$$\begin{aligned} \text{min.} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \\ \text{s.t.} \quad & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{aligned}$$

性質

P と L($\boldsymbol{\lambda}$) が最適解を持つ $\Rightarrow$ P の最適値 $\geq$ L($\boldsymbol{\lambda}$) の最適値

ラグランジュ緩和：性質 (続)

整数計画問題 P $\longrightarrow$ ラグランジュ緩和 L($\boldsymbol{\lambda}$)

$$\begin{array}{ll} \text{min.} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{s.t.} & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & A_2\boldsymbol{x} = \boldsymbol{b}_2, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{array}$$

$$\begin{array}{ll} \text{min.} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \\ \text{s.t.} & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{array}$$

緩和問題に求められる性質

- 解きたい問題 P の許容領域 $\subseteq$ 緩和問題の許容領域
- 解きたい問題 P の最適値 $\geq$ 緩和問題の最適値
- 解きたい問題 P よりも 緩和問題の方が解きやすい

ラグランジュ双対問題

整数計画問題 P → ラグランジュ緩和 L($\boldsymbol{\lambda}$)

$$\begin{array}{ll} \text{min.} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{s.t.} & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & A_2\boldsymbol{x} = \boldsymbol{b}_2, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{array}$$

$$\begin{array}{ll} \text{min.} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \\ \text{s.t.} & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{array}$$

L($\boldsymbol{\lambda}$) の最適値として得られる
P の最適値の下界として
最大のものを求める問題

ラグランジュ双対 LD

$$\max_{\boldsymbol{\lambda}} \quad \begin{array}{ll} \min\limits_{\boldsymbol{x}} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \\ \text{s.t.} & A_1\boldsymbol{x} \geq \boldsymbol{b}_1, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{array}$$

(Lagrangian dual)

最強ラグランジュ緩和

整数計画問題 P $\longrightarrow$ 最強ラグランジュ緩和 L$(\boldsymbol{\lambda}^*)$

$$\begin{aligned} \text{min.} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{s.t.} \quad & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & A_2\boldsymbol{x} = \boldsymbol{b}_2, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{aligned}$$

$$\begin{aligned} \text{min.} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{*\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \\ \text{s.t.} \quad & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{aligned}$$

L$(\boldsymbol{\lambda})$ の最適値として得られる
P の最適値の下界として
最大のものを求める問題

$\boldsymbol{\lambda}^* =$ LD の最適解

ラグランジュ双対 LD

$$\max_{\boldsymbol{\lambda}} \quad \begin{aligned} \min_{\boldsymbol{x}} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \\ \text{s.t.} \quad & A_1\boldsymbol{x} \geq \boldsymbol{b}_1, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \\ & \boldsymbol{x} \in \mathbb{Z}^n \end{aligned}$$

(Lagrangian dual)

ラグランジュ緩和と 1-木

前回の内容

巡回セールスマン問題のラグランジュ緩和として
よいものを作る

これは Held, Karp ('70, '71) によるアイディア

よさ の基準

- 得られる下界が大きい
- 効率よく解ける
- 運がよいと,
  巡回路が得られる

解きたい問題 P

$$
\begin{aligned}
\text{min.} \quad & \sum_{e \in E} d(e) x_e & \\
\text{s.t.} \quad & \sum_{e \in \delta(i)} x_e = 2 & \forall\, i \\
& \sum_{e \subseteq S} x_e \leq |S| - 1 & \forall\, S \\
& x_e \in \{0, 1\} & \forall\, e
\end{aligned}
$$

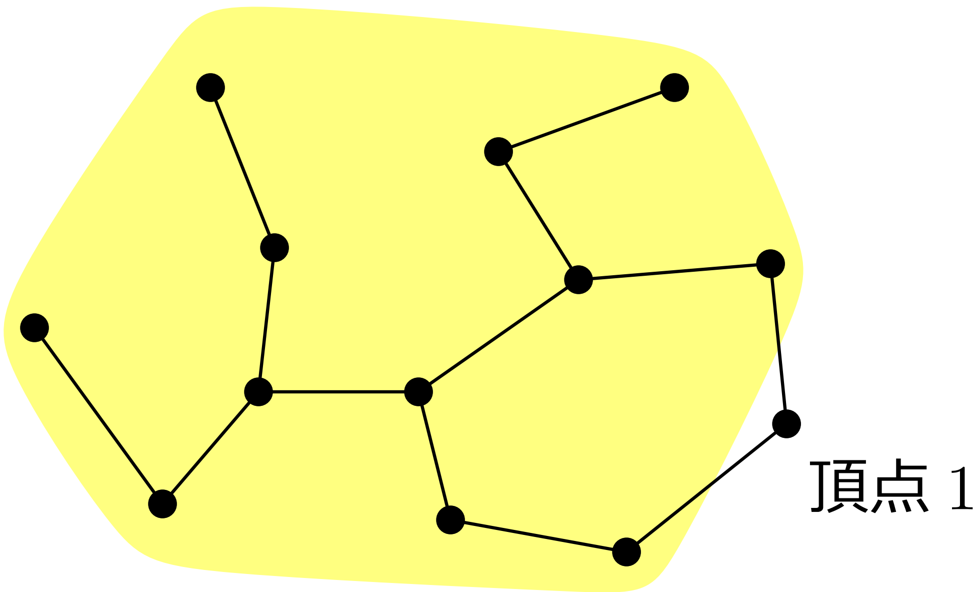
最小重み 1-木問題のアルゴリズム

最小重み 1-木問題のアルゴリズム

1. $\{2, 3, \ldots, n\}$ 上の最小重み全域木を求める
2. 頂点 1 に接続する辺の中で重みの小さい方から 2 辺を求める
3. それらの合併を出力する

Kruskal 法などで
効率よく実行できる

これは多項式時間アルゴリズム



したがって，
いま作ったラグランジュ緩和は
効率よく解ける

最強ラグランジュ緩和を求めたい

ラグランジュ緩和 $L(\boldsymbol{\lambda})$ $\xrightarrow[\text{計算できる}]{}$ 最適解 (最小重み 1-木)

効率よく
計算できる

欲しいもの

最強ラグランジュ緩和 $L(\boldsymbol{\lambda}^*)$

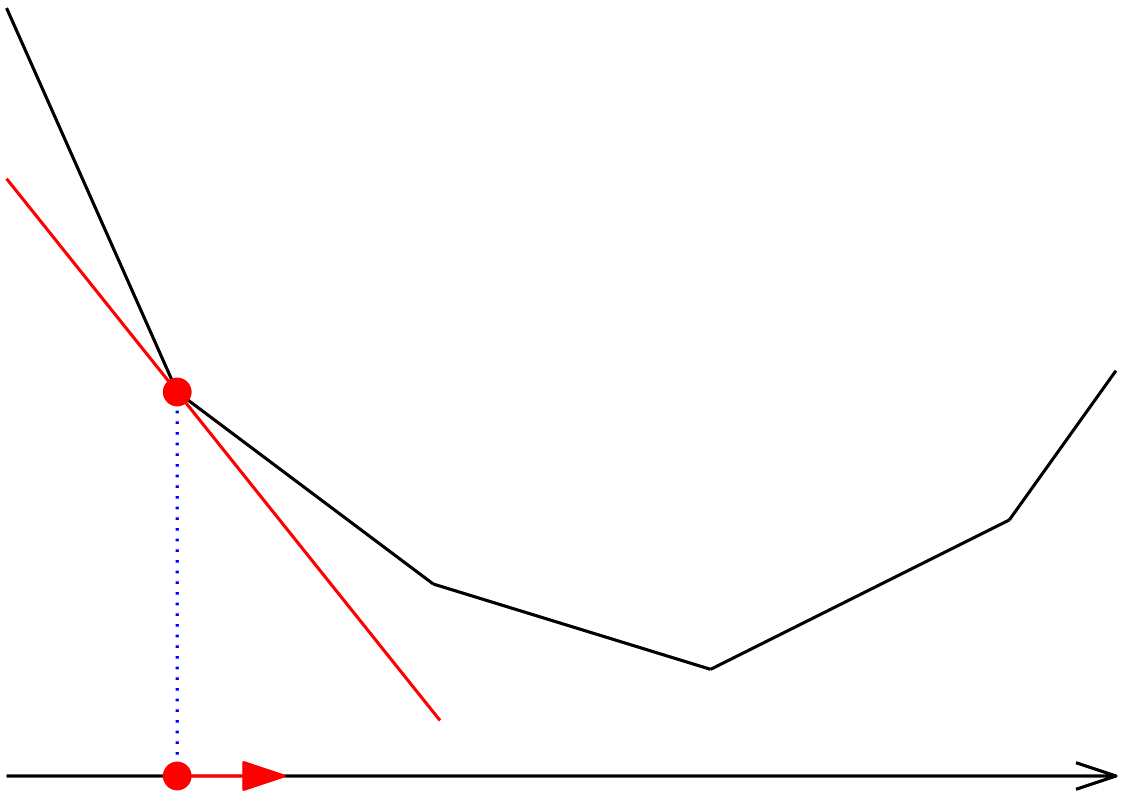
$\boldsymbol{\lambda}^*$ = ラグランジュ緩和 $L(\boldsymbol{\lambda})$ の最適値を最大にする $\boldsymbol{\lambda}$

↓

どのように $\boldsymbol{\lambda}^*$ を求めるのか → 今回 (劣勾配法)

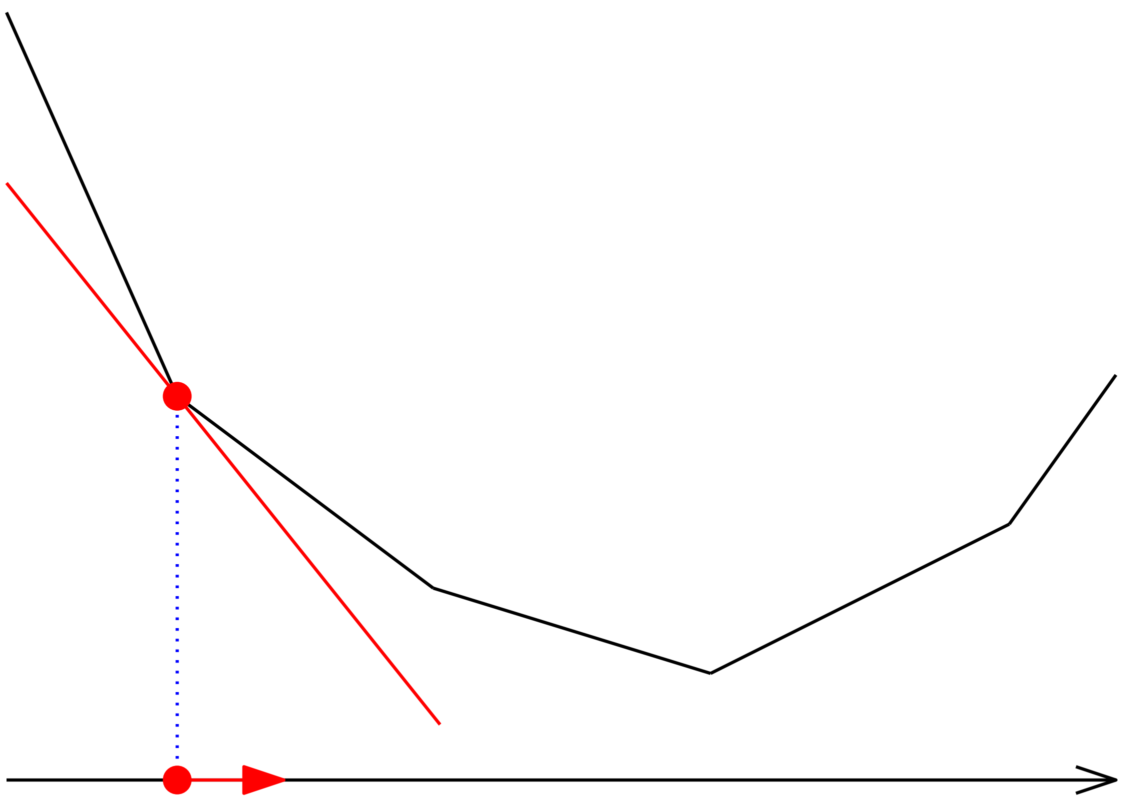
今日の内容

- ラグランジュ緩和：復習
- ラグランジュ緩和の性質
- 劣微分と劣勾配
- 劣勾配法



今日の内容

- ラグランジュ緩和：復習
- ラグランジュ緩和の性質
- 劣微分と劣勾配
- 劣勾配法



ラグランジュ緩和の最適値関数

ラグランジュ緩和 L($\boldsymbol{\lambda}$)

$$
\begin{aligned}
\text{min.}\quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \\
\text{s.t.}\quad & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\
& \boldsymbol{x} \geq \boldsymbol{0}, \\
& \boldsymbol{x} \in \mathbb{Z}^n
\end{aligned}
$$

最適値は $\boldsymbol{\lambda}$ に依存する

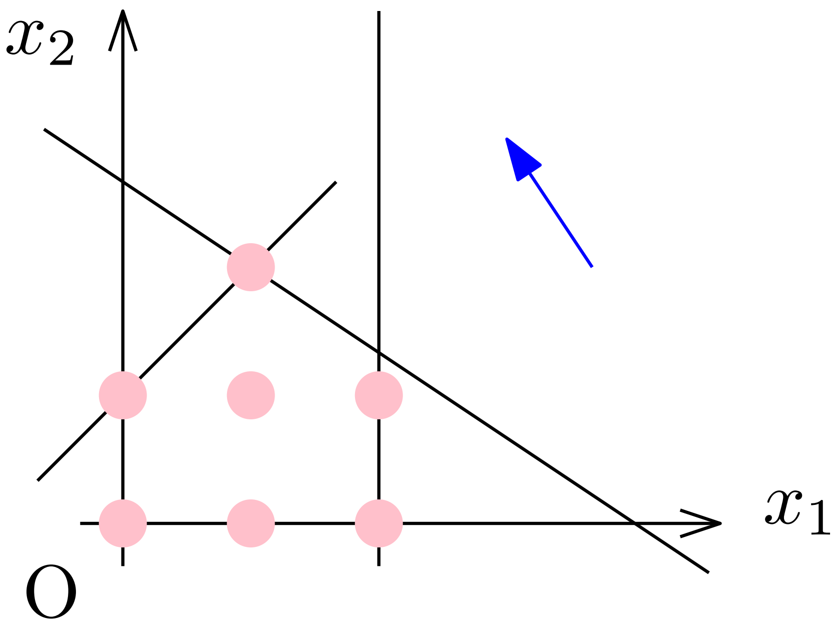
$= g(\boldsymbol{\lambda})$ とする

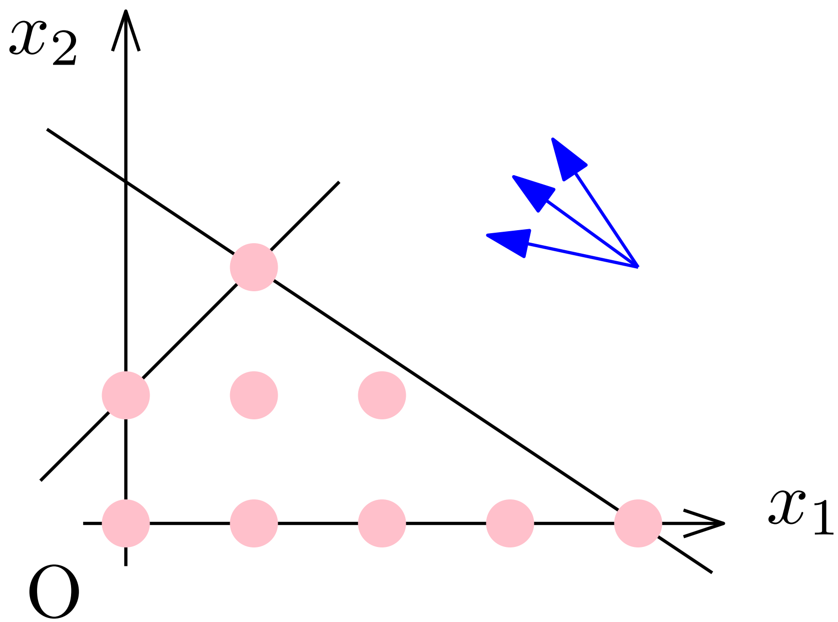
ラグランジュ緩和：例

整数計画問題 P → ラグランジュ緩和 L(λ)

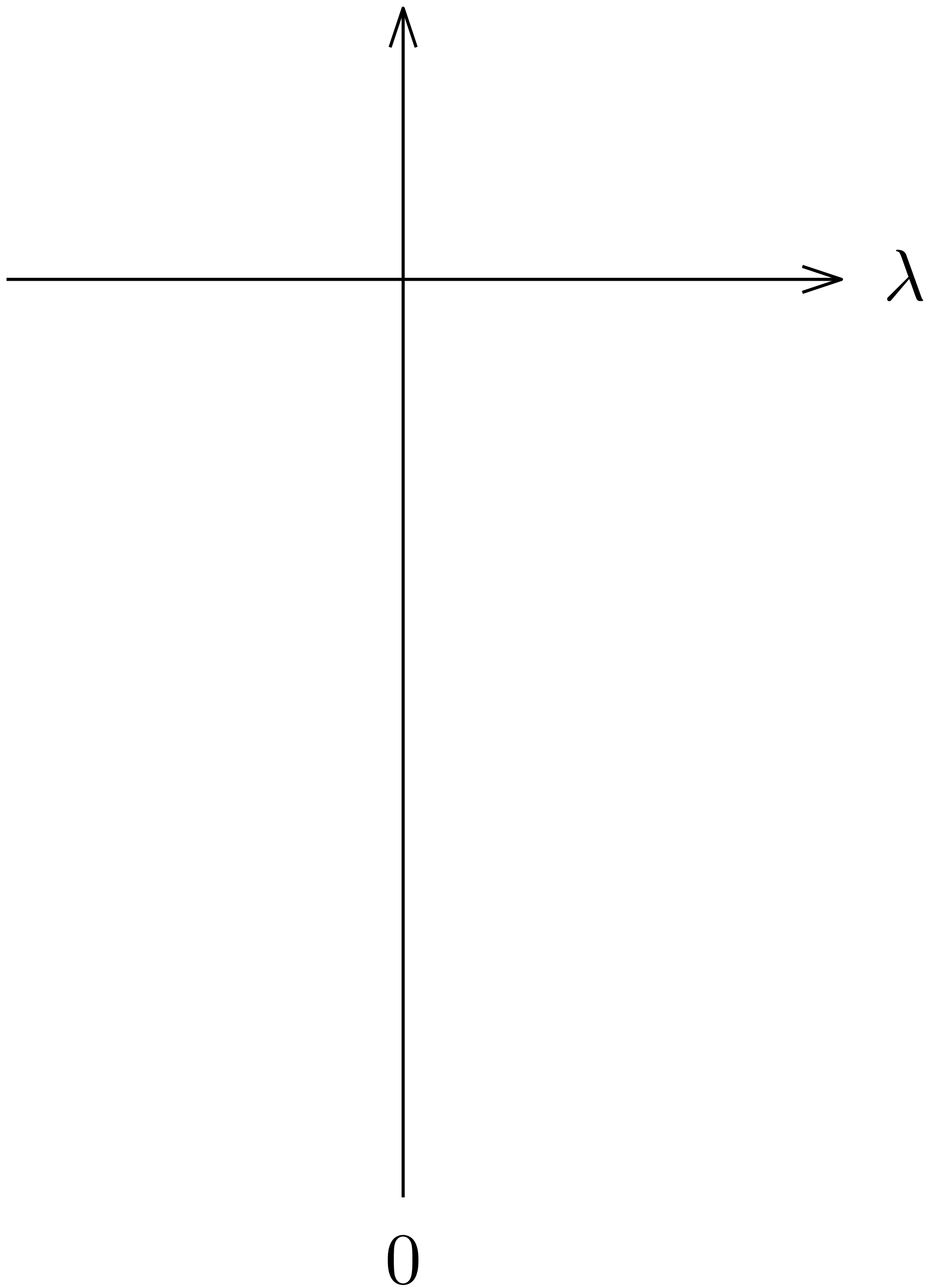
$$
\begin{aligned}
\text{min.}\ & 2x_1 - 3x_2 \\
\text{s.t.}\ & -x_1 + x_2 \le 1, \\
& x_1 \le 2, \\
& 4x_1 + 5x_2 \le 20, \\
& x_1, x_2 \ge 0, \\
& x_1, x_2 \in \mathbb{Z}
\end{aligned}
$$

$$
\begin{aligned}
\text{min.}\ & 2x_1 - 3x_2 + \lambda(x_1 - 2) \\
\text{s.t.}\ & -x_1 + x_2 \le 1, \\
& 4x_1 + 5x_2 \le 20, \\
& x_1, x_2 \ge 0, \\
& x_1, x_2 \in \mathbb{Z}
\end{aligned}
$$



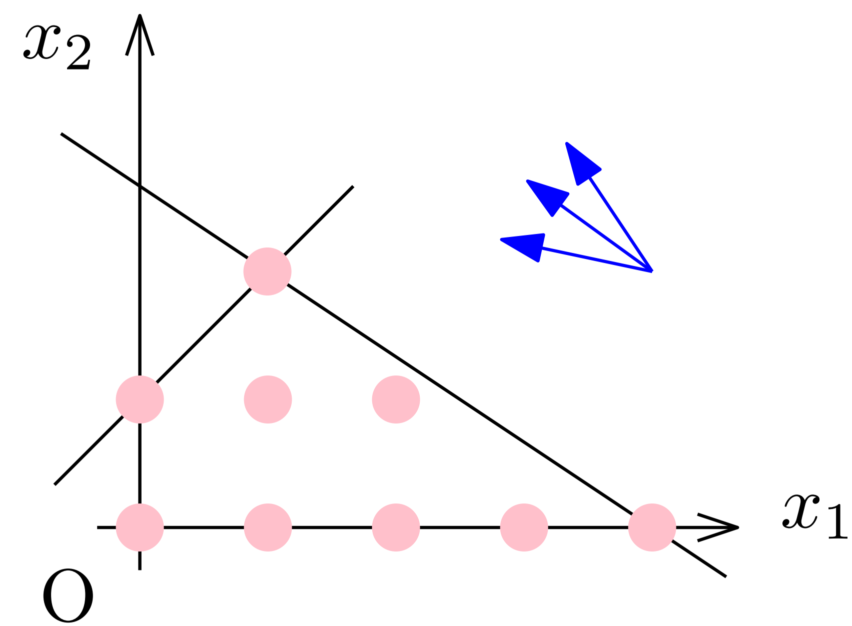


最適値関数：例

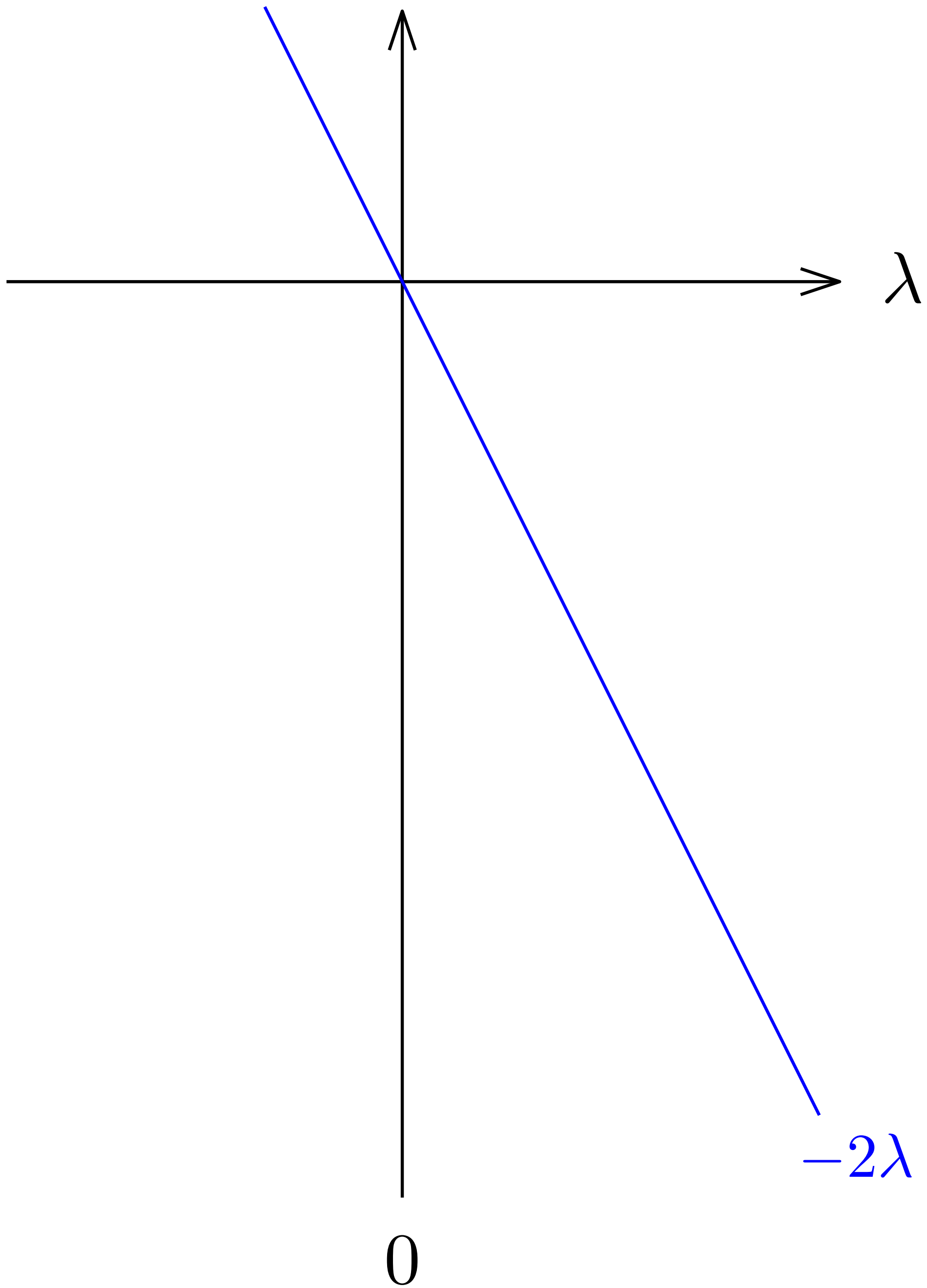


ラグランジュ緩和 L(λ)

$$\begin{aligned} \text{min.}\ & 2x_1 - 3x_2 + \lambda(x_1 - 2) \\ & = (2+\lambda)x_1 - 3x_2 - 2\lambda \\ \text{s.t.}\ & -x_1 + x_2 \le 1, \\ & 4x_1 + 5x_2 \le 20, \\ & x_1, x_2 \ge 0, \\ & x_1, x_2 \in \mathbb{Z} \end{aligned}$$

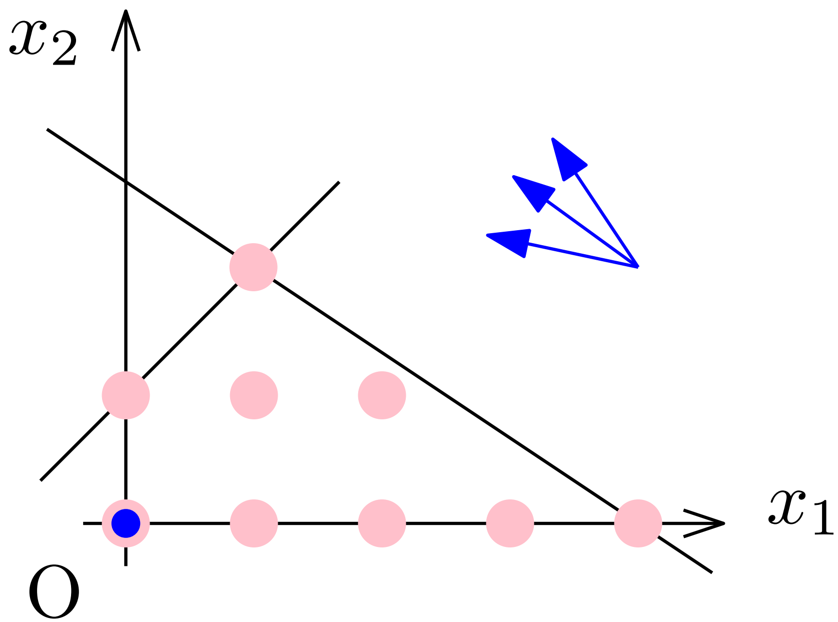


最適値関数：例

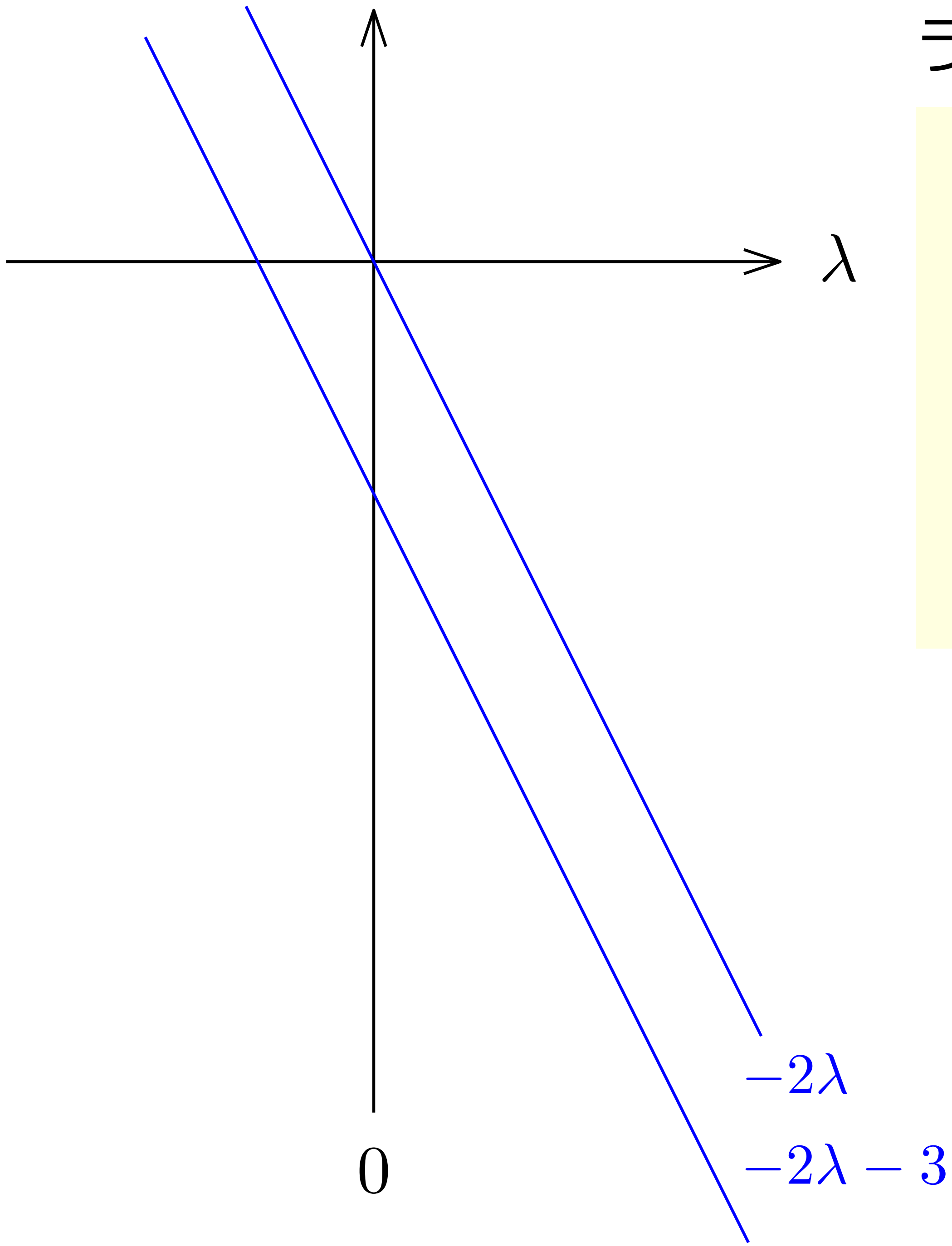


ラグランジュ緩和 L(λ)

$$\begin{aligned} \text{min.}\ & 2x_1 - 3x_2 + \lambda(x_1 - 2) \\ & = (2+\lambda)x_1 - 3x_2 - 2\lambda \\ \text{s.t.}\ & -x_1 + x_2 \le 1, \\ & 4x_1 + 5x_2 \le 20, \\ & x_1, x_2 \ge 0, \\ & x_1, x_2 \in \mathbb{Z} \end{aligned}$$

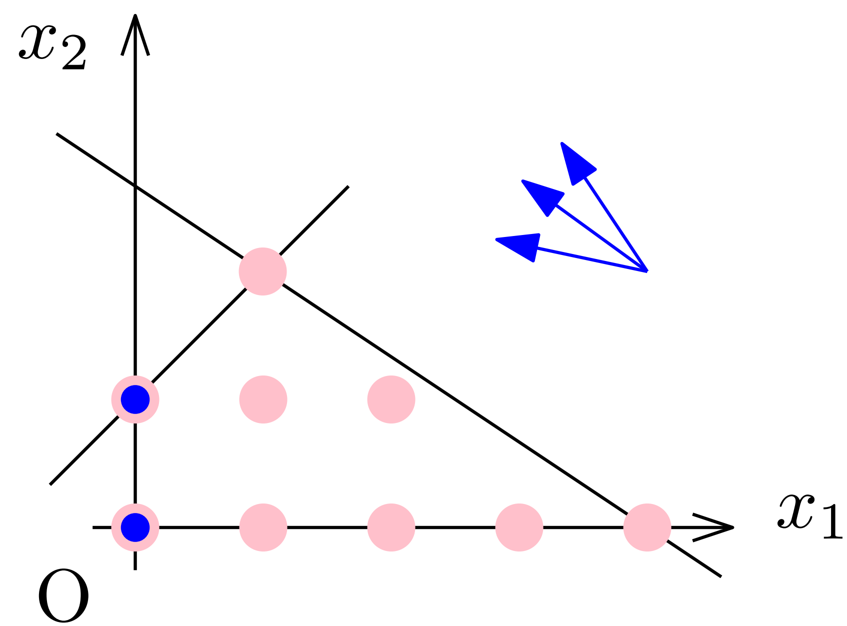


最適値関数：例

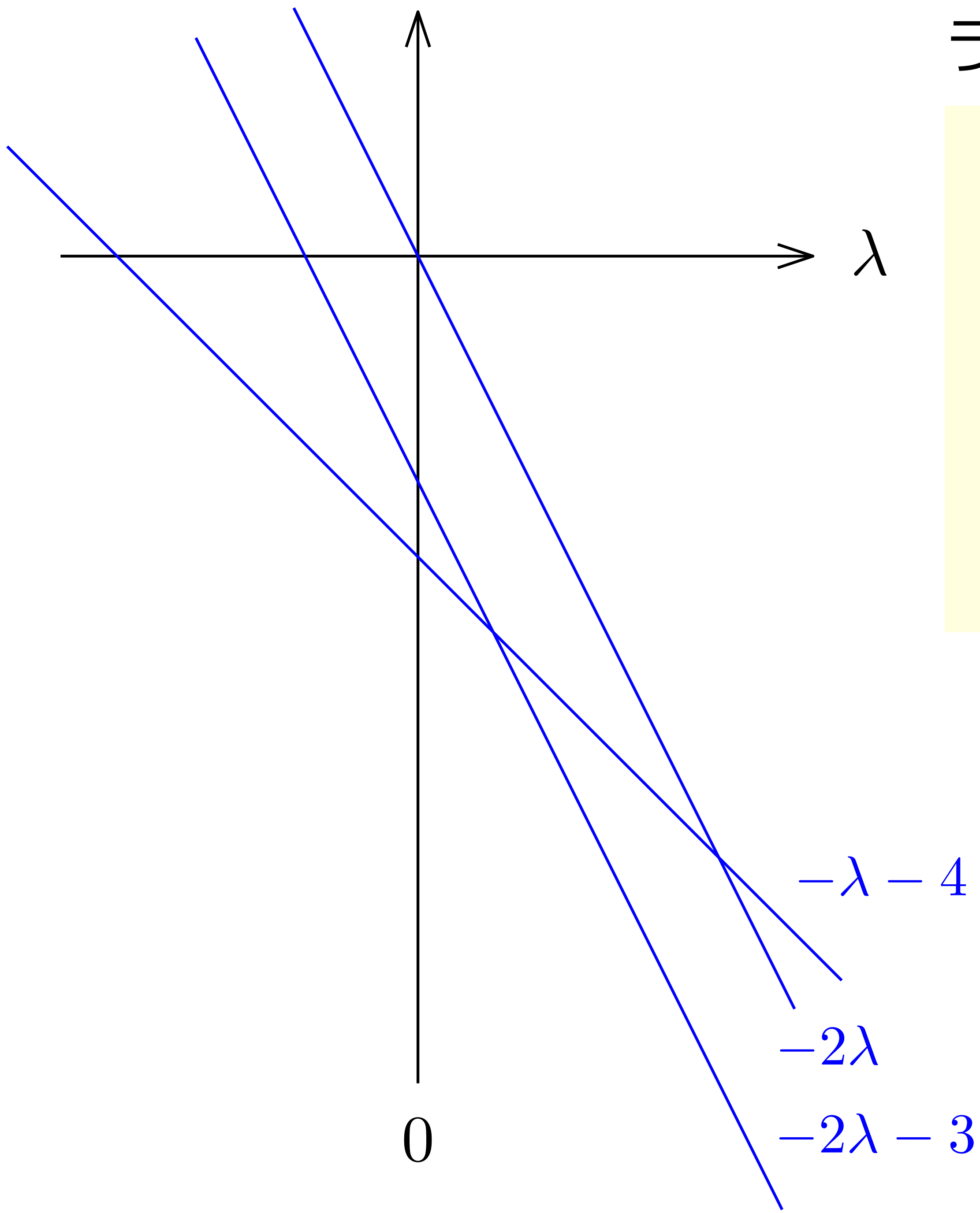


ラグランジュ緩和 L(λ)

$$
\begin{aligned}
\text{min.}\ & 2x_1 - 3x_2 + \lambda(x_1 - 2) \\
& = (2+\lambda)x_1 - 3x_2 - 2\lambda \\
\text{s.t.}\ & -x_1 + x_2 \le 1, \\
& 4x_1 + 5x_2 \le 20, \\
& x_1, x_2 \ge 0, \\
& x_1, x_2 \in \mathbb{Z}
\end{aligned}
$$

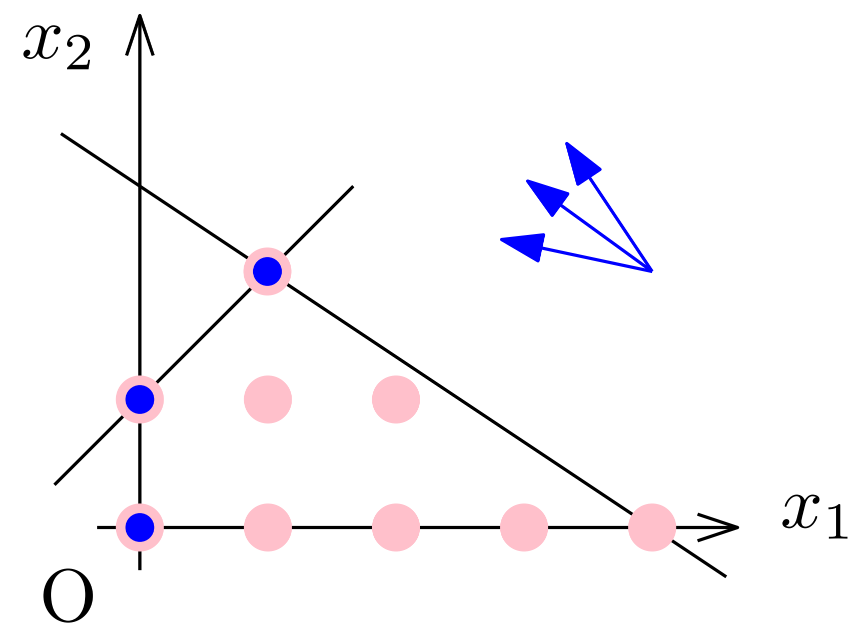


最適値関数：例

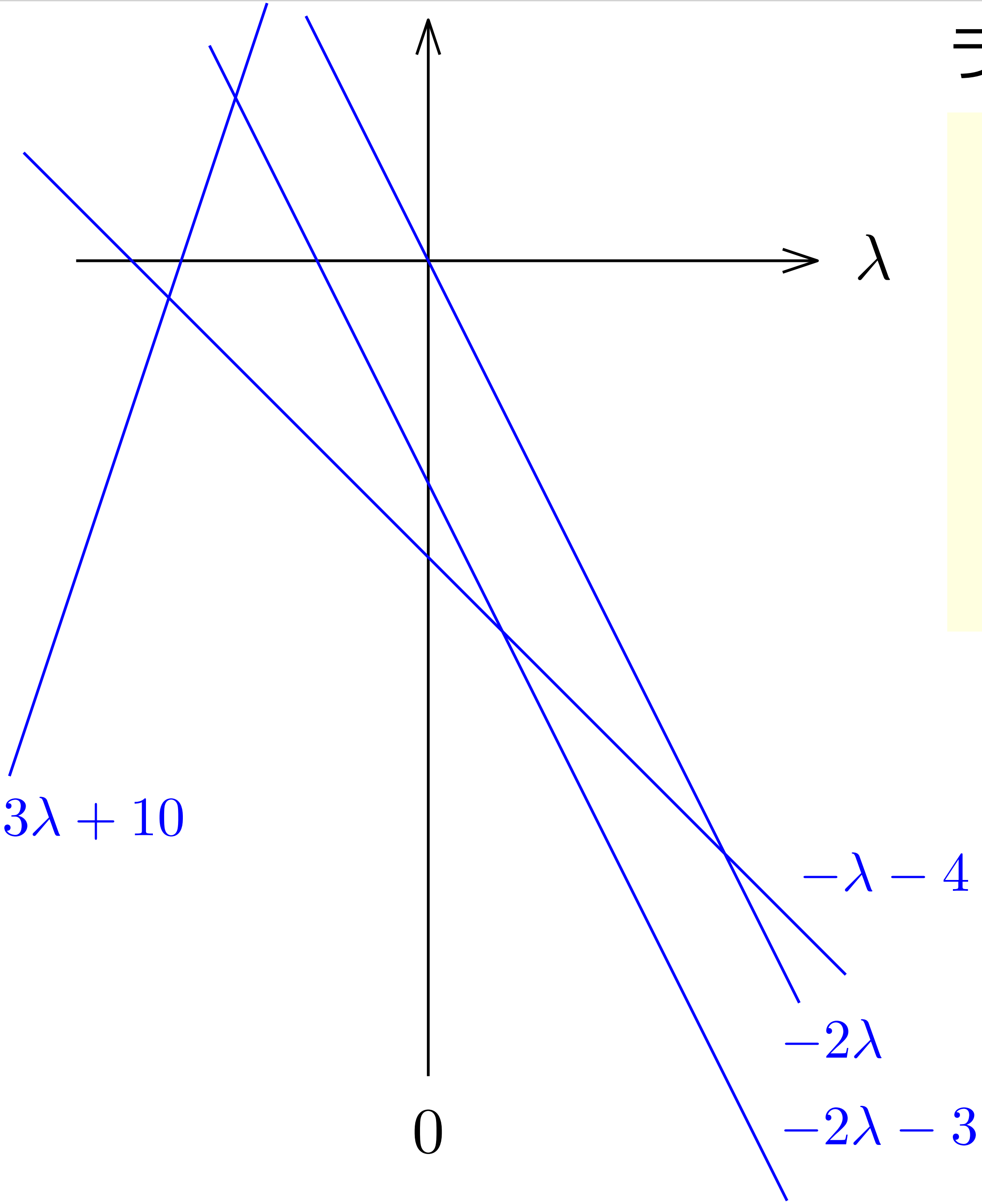


ラグランジュ緩和 L(λ)

$$\begin{aligned}
\text{min.}\ & 2x_1 - 3x_2 + \lambda(x_1 - 2) \\
& = (2+\lambda)x_1 - 3x_2 - 2\lambda \\
\text{s.t.}\ & -x_1 + x_2 \le 1, \\
& 4x_1 + 5x_2 \le 20, \\
& x_1, x_2 \ge 0, \\
& x_1, x_2 \in \mathbb{Z}
\end{aligned}$$

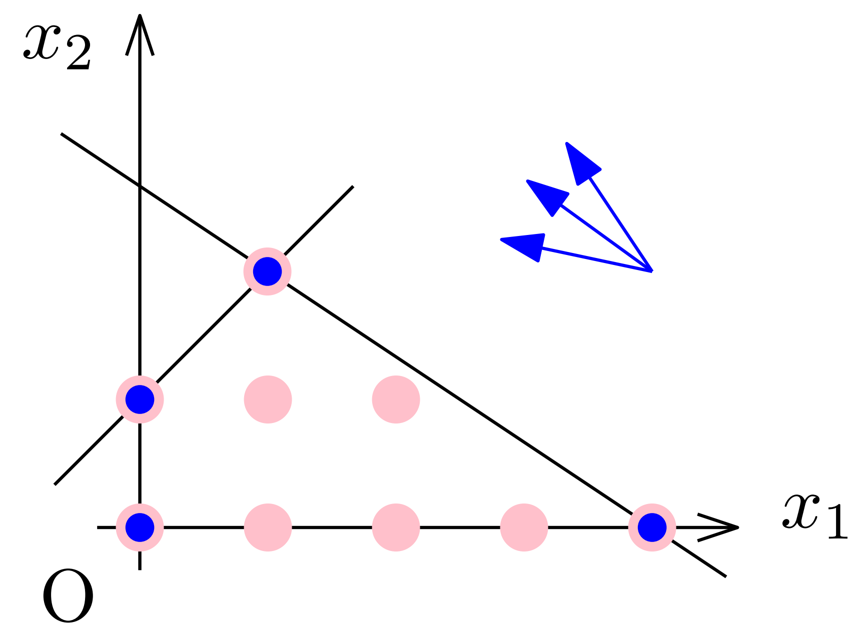


最適値関数：例

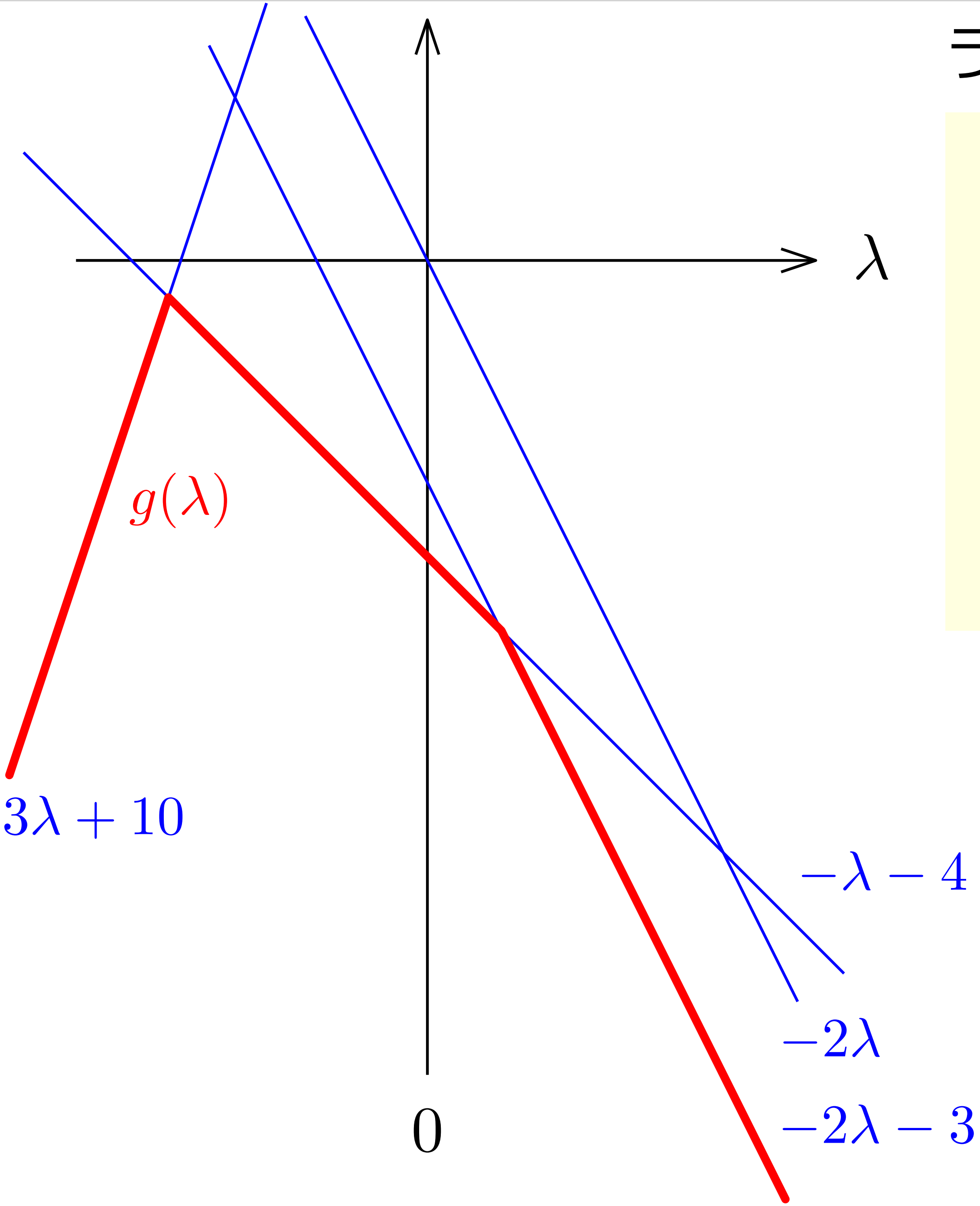


ラグランジュ緩和 L(λ)

$$
\begin{aligned}
\text{min.}\ & 2x_1 - 3x_2 + \lambda(x_1 - 2) \\
& \quad = (2+\lambda)x_1 - 3x_2 - 2\lambda \\
\text{s.t.}\ & -x_1 + x_2 \le 1, \\
& 4x_1 + 5x_2 \le 20, \\
& x_1, x_2 \ge 0, \\
& x_1, x_2 \in \mathbb{Z}
\end{aligned}
$$

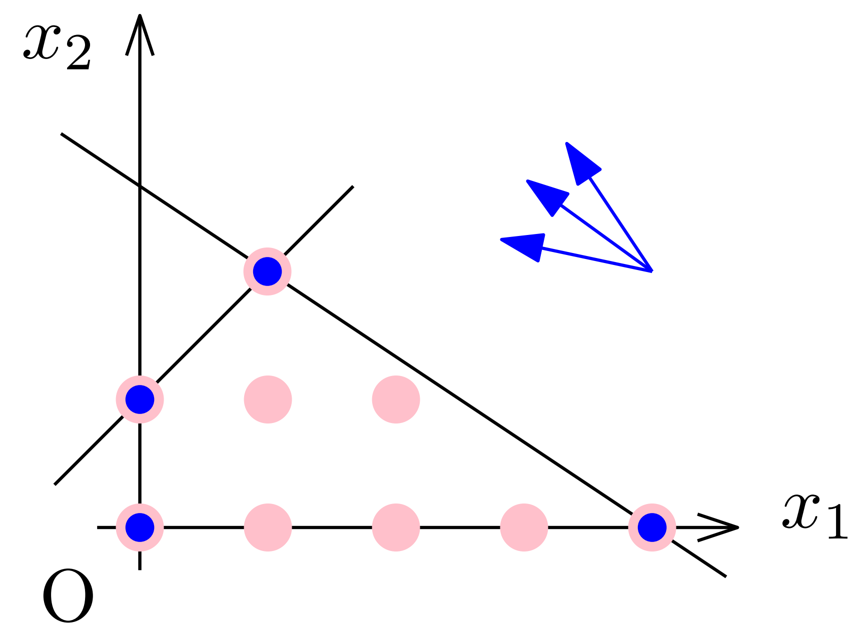


最適値関数：例



ラグランジュ緩和 L(λ)

$$\begin{aligned} \text{min.}\ & 2x_1 - 3x_2 + \lambda(x_1 - 2) \\ & = (2+\lambda)x_1 - 3x_2 - 2\lambda \\ \text{s.t.}\ & -x_1 + x_2 \le 1, \\ & 4x_1 + 5x_2 \le 20, \\ & x_1, x_2 \ge 0, \\ & x_1, x_2 \in \mathbb{Z} \end{aligned}$$



最適値関数は凹関数

ラグランジュ緩和 L($\boldsymbol{\lambda}$)

$$
\begin{aligned}
\text{min.}\quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \\
\text{s.t.}\quad & A_1\boldsymbol{x} = \boldsymbol{b}_1, \\
& \boldsymbol{x} \geq \boldsymbol{0}, \\
& \boldsymbol{x} \in \mathbb{Z}^n
\end{aligned}
$$

最適値は $\boldsymbol{\lambda}$ に依存する

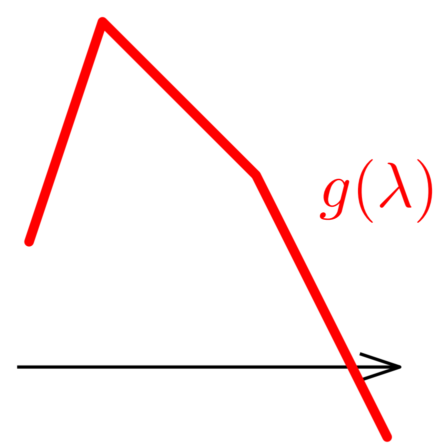
$= g(\boldsymbol{\lambda})$ とする

$g \colon \mathbb{R}^{m_2} \to \mathbb{R}$

性質：最適値関数は凹関数

最適値関数 g は凹関数

凹関数の定義は次のページで



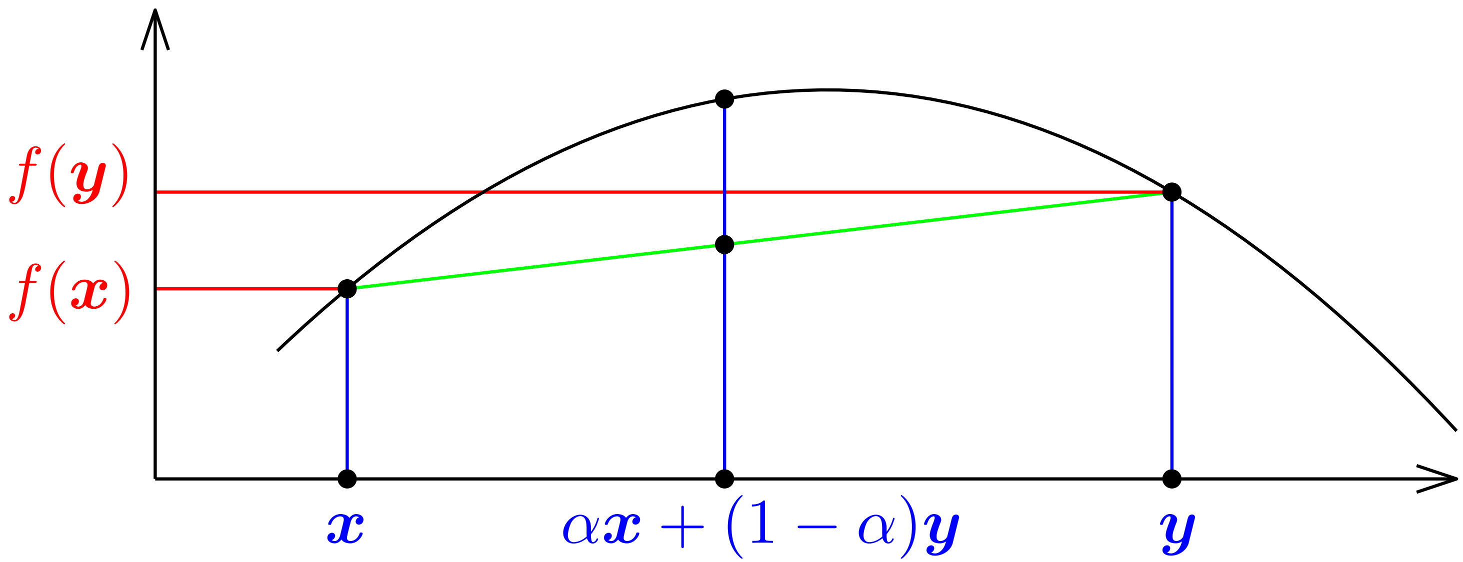
定義：凹関数

定義：凹関数 (concave function)

関数 $f\colon \mathbb{R}^n \to \mathbb{R}$ が **凹関数** であるとは，次を満たすこと

任意の $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ と $\alpha \in [0,1]$ に対して

$$f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \geq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y})$$



f の始域は凸集合でもよい

最適値関数は凹関数：証明 (1)

証明： 任意の $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^{m_2}$， $\alpha \in [0, 1]$ を考える

$\boxed{\text{目標}}$ ：$g(\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu}) \geq \alpha g(\boldsymbol{\lambda}) + (1-\alpha) g(\boldsymbol{\mu})$

最適値関数 g の定義より，

$$g(\boldsymbol{\lambda}) = \min\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \mid A_1\boldsymbol{x} = \boldsymbol{b}_1, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathbb{Z}^n\}$$

→ 最適解を $\boldsymbol{x}_{\boldsymbol{\lambda}}$ とする

$$g(\boldsymbol{\mu}) = \min\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \mid A_1\boldsymbol{x} = \boldsymbol{b}_1, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathbb{Z}^n\}$$

→ 最適解を $\boldsymbol{x}_{\boldsymbol{\mu}}$ とする

最適値関数は凹関数：証明 (1)

証明： 任意の $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^{m_2}$， $\alpha \in [0, 1]$ を考える

$\boxed{\text{目標}}$: $g(\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu}) \geq \alpha g(\boldsymbol{\lambda}) + (1-\alpha)g(\boldsymbol{\mu})$

最適値関数 g の定義より，

$$g(\boldsymbol{\lambda}) = \min\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \mid A_1\boldsymbol{x} = \boldsymbol{b}_1, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathbb{Z}^n\}$$

最適解を $\boldsymbol{x}_{\boldsymbol{\lambda}}$ とする

$$g(\boldsymbol{\mu}) = \min\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \mid A_1\boldsymbol{x} = \boldsymbol{b}_1, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathbb{Z}^n\}$$

最適解を $\boldsymbol{x}_{\boldsymbol{\mu}}$ とする

したがって，

$$g(\boldsymbol{\lambda}) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\lambda}} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\lambda}})$$

$$g(\boldsymbol{\mu}) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\mu}} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\mu}})$$

最適値関数は凹関数：証明 (2)

証明： 任意の $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^m$, $\alpha \in [0,1]$ を考える

$\boxed{\text{目標}}$: $g(\underbrace{\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu}}_{=\boldsymbol{\nu}\text{ とする}}) \geq \alpha g(\boldsymbol{\lambda}) + (1-\alpha) g(\boldsymbol{\mu})$

最適値関数 g の定義より,

$$g(\boldsymbol{\nu}) = \min\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\nu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \mid A_1\boldsymbol{x} = \boldsymbol{b}_1, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathbb{Z}^n\}$$

→ 最適解を $\boldsymbol{x}_{\boldsymbol{\nu}}$ とする

$$\begin{aligned} g(\boldsymbol{\nu}) &= \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + \boldsymbol{\nu}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}}) \\ &= (\alpha + (1-\alpha))\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + (\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu})^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}}) \\ &= \alpha(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}})) \\ &\quad + (1-\alpha)(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}})) \end{aligned}$$

最適値関数は凹関数：証明 (2)

証明： 任意の $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^m$， $\alpha \in [0,1]$ を考える

$\boxed{\text{目標}}$：$g(\underbrace{\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu}}_{=\boldsymbol{\nu}\text{ とする}}) \geq \alpha g(\boldsymbol{\lambda}) + (1-\alpha) g(\boldsymbol{\mu})$

最適値関数 g の定義より，

$$g(\boldsymbol{\nu}) = \min\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\nu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \mid A_1\boldsymbol{x} = \boldsymbol{b}_1, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathbb{Z}^n\}$$

→ 最適解を $\boldsymbol{x}_{\boldsymbol{\nu}}$ とする

$$\begin{aligned}
g(\boldsymbol{\nu}) &= \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + \boldsymbol{\nu}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}}) \\
&= (\alpha + (1-\alpha))\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + (\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu})^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}}) \\
&= \alpha(\underbrace{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}})}_{\geq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\lambda}} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\lambda}}) = g(\boldsymbol{\lambda})}) \\
&\quad + (1-\alpha)(\underbrace{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}})}_{\geq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\mu}} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\mu}}) = g(\boldsymbol{\mu})})
\end{aligned}$$

最適値関数は凹関数：証明 (2)

証明： 任意の $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathbb{R}^m$, $\alpha \in [0, 1]$ を考える

$\boxed{\text{目標}}$: $g(\underbrace{\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu}}_{=\boldsymbol{\nu}\text{ とする}}) \geq \alpha g(\boldsymbol{\lambda}) + (1-\alpha) g(\boldsymbol{\mu})$

最適値関数 g の定義より，

$$g(\boldsymbol{\nu}) = \min\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\nu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \mid A_1\boldsymbol{x} = \boldsymbol{b}_1, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathbb{Z}^n\}$$

最適解を $\boldsymbol{x}_{\boldsymbol{\nu}}$ とする

$$\begin{aligned}
g(\boldsymbol{\nu}) &= \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + \boldsymbol{\nu}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}}) \\
&= (\alpha + (1-\alpha))\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + (\alpha\boldsymbol{\lambda} + (1-\alpha)\boldsymbol{\mu})^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}}) \\
&= \alpha(\underbrace{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}})}_{\geq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\lambda}} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\lambda}}) = g(\boldsymbol{\lambda})}) \\
&\quad + (1-\alpha)(\underbrace{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\nu}} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\nu}})}_{\geq \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\mu}} + \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\mu}}) = g(\boldsymbol{\mu})}) \\
&\geq \alpha g(\boldsymbol{\lambda}) + (1-\alpha) g(\boldsymbol{\mu})
\end{aligned}$$

$\square$

凹関数の最大化

はじめの目標

$\boldsymbol{\lambda}^* =$ ラグランジュ緩和 $\mathsf{L}(\boldsymbol{\lambda})$ の最適値を最大にする $\boldsymbol{\lambda}$ を求めること

凹関数の最大化

はじめの目標

$\boldsymbol{\lambda}^* = \underbrace{\text{ラグランジュ緩和 } \mathsf{L}(\boldsymbol{\lambda}) \text{ の最適値}}_{= g(\boldsymbol{\lambda})}$ を最大にする $\boldsymbol{\lambda}$ を求めること

証明した性質：g は凹関数

$\downarrow$

できれば十分であること

凹関数の最大化

使用できるアルゴリズム，有効なアルゴリズムは最大化したい凹関数の性質に依存する

微分不可能凹関数の最大化

できれば十分であること

凹関数の最大化

使用できるアルゴリズム，有効なアルゴリズムは
最大化したい凹関数の性質に依存する

ラグランジュ緩和の最適値関数 g の性質

- 凹関数
- 微分可能であるとは限らない (non-smooth)

今から紹介する **劣勾配法** (subgradient method) は
微分不可能な凹関数の最大化を行う伝統的なアルゴリズム

劣勾配法は，普通「凸関数の最小化」に対して記述される

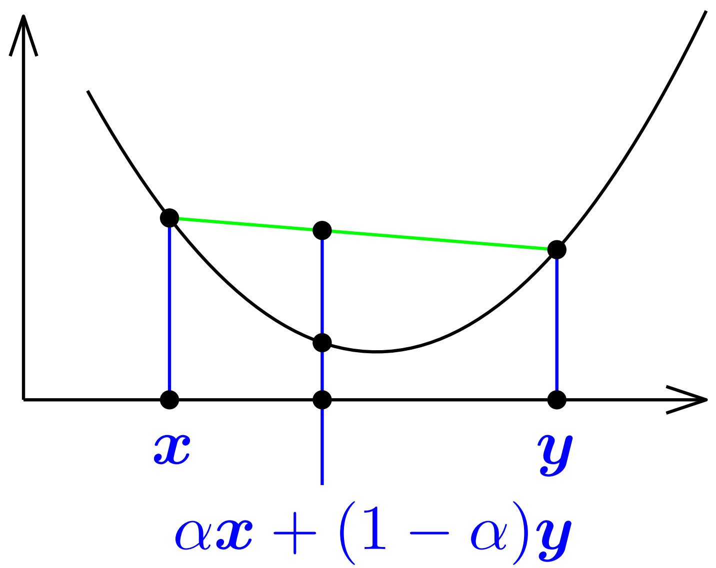
定義：凸関数 (convex function)

関数 $f\colon \mathbb{R}^n \to \mathbb{R}$ が **凸関数** であるとは，次を満たすこと

任意の $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ と $\alpha \in [0,1]$ に対して

$$f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \le \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y})$$

凹関数では「$\geq$」だった



f の始域は凸集合でもよい

$\boxed{\text{性質}}$ f が凹関数 $\Leftrightarrow$ $-f$ が凸関数

つまり，凹関数の最大化は

凸関数の最小化と変わらない

凸関数の性質：極小解 = 最小解

性質：凸関数にて極小解は最小解

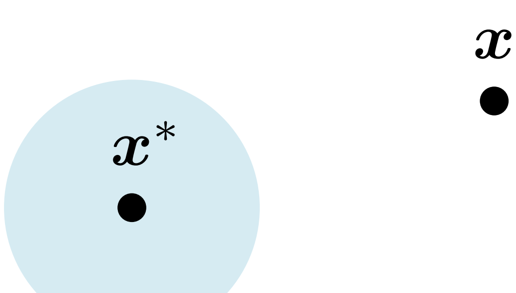
$\boldsymbol{x}^* \in \mathbb{R}^n$ が次を満たすとする

$\boldsymbol{x}^*$ に近い任意の $\tilde{\boldsymbol{x}}$ に対して，$f(\boldsymbol{x}^*) \leq f(\tilde{\boldsymbol{x}})$

このとき，次が成り立つ

任意の $\boldsymbol{x} \in \mathbb{R}^n$ に対して，$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$

<u>証明</u>：ある $\boldsymbol{x} \in \mathbb{R}^n$ に対して，$f(\boldsymbol{x}) < f(\boldsymbol{x}^*)$ であると仮定する



凸関数の性質：極小解 = 最小解

性質：凸関数にて極小解は最小解

$\boldsymbol{x}^* \in \mathbb{R}^n$ が次を満たすとする

　$\boldsymbol{x}^*$ に近い任意の $\tilde{\boldsymbol{x}}$ に対して，$f(\boldsymbol{x}^*) \leq f(\tilde{\boldsymbol{x}})$

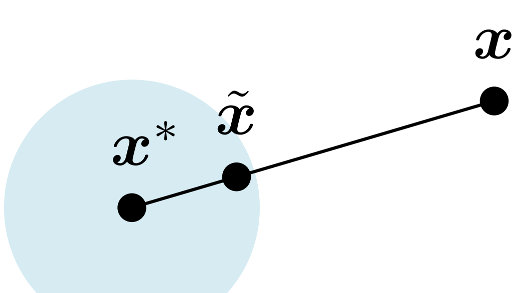
このとき，次が成り立つ

　任意の $\boldsymbol{x} \in \mathbb{R}^n$ に対して，$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$

証明：ある $\boldsymbol{x} \in \mathbb{R}^n$ に対して，$f(\boldsymbol{x}) < f(\boldsymbol{x}^*)$ であると仮定する

$\boldsymbol{x}$ と $\boldsymbol{x}^*$ の内分点として，$\boldsymbol{x}^*$ に近い $\tilde{\boldsymbol{x}}$ を取れる

ある $\alpha \in (0,1)$ を用いて，$\tilde{\boldsymbol{x}} = \alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{x}^*$ とする



凸関数の性質：極小解 ＝ 最小解

性質：凸関数にて極小解は最小解

$\boldsymbol{x}^* \in \mathbb{R}^n$ が次を満たすとする

$\boldsymbol{x}^*$ に近い任意の $\tilde{\boldsymbol{x}}$ に対して，$f(\boldsymbol{x}^*) \leq f(\tilde{\boldsymbol{x}})$

このとき，次が成り立つ

任意の $\boldsymbol{x} \in \mathbb{R}^n$ に対して，$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$

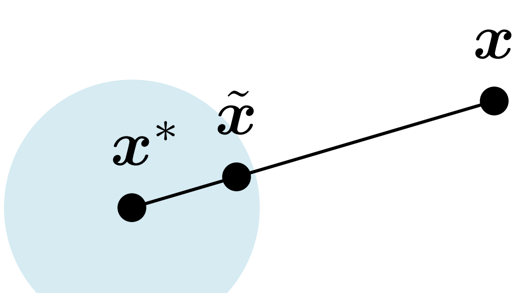
証明：ある $\boldsymbol{x} \in \mathbb{R}^n$ に対して，$f(\boldsymbol{x}) < f(\boldsymbol{x}^*)$ であると仮定する

$\boldsymbol{x}$ と $\boldsymbol{x}^*$ の内分点として，$\boldsymbol{x}^*$ に近い $\tilde{\boldsymbol{x}}$ を取れる

ある $\alpha \in (0,1)$ を用いて，$\tilde{\boldsymbol{x}} = \alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{x}^*$ とする

$$f(\boldsymbol{x}^*) \leq f(\tilde{\boldsymbol{x}}) = f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{x}^*) \leq \alpha f(\boldsymbol{x}) + (1-\alpha)f(\boldsymbol{x}^*)$$
$$< \alpha f(\boldsymbol{x}^*) + (1-\alpha)f(\boldsymbol{x}^*) = f(\boldsymbol{x}^*)$$

$\lightning$



凸関数の最小化：微分可能である場合

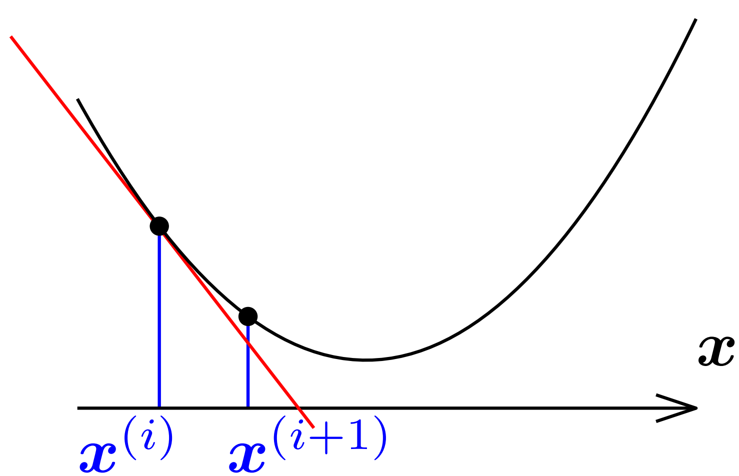
凸関数 $f\colon \mathbb{R}^n \to \mathbb{R}$ が微分可能であるならば ...

アルゴリズム：勾配法 (gradient method)

1. 初期点 $\boldsymbol{x}^{(0)}$ を適当に選ぶ
2. $i = 0, 1, \ldots$ で収束するまで以下を繰り返す
   $$\boldsymbol{x}^{(i+1)} := \boldsymbol{x}^{(i)} - \alpha_i \underbrace{\nabla f(\boldsymbol{x}^{(i)})}_{\boldsymbol{x}^{(i)}\text{ における } f \text{ の勾配}}$$
   ($\alpha_i > 0$ はうまく選んだ実数)

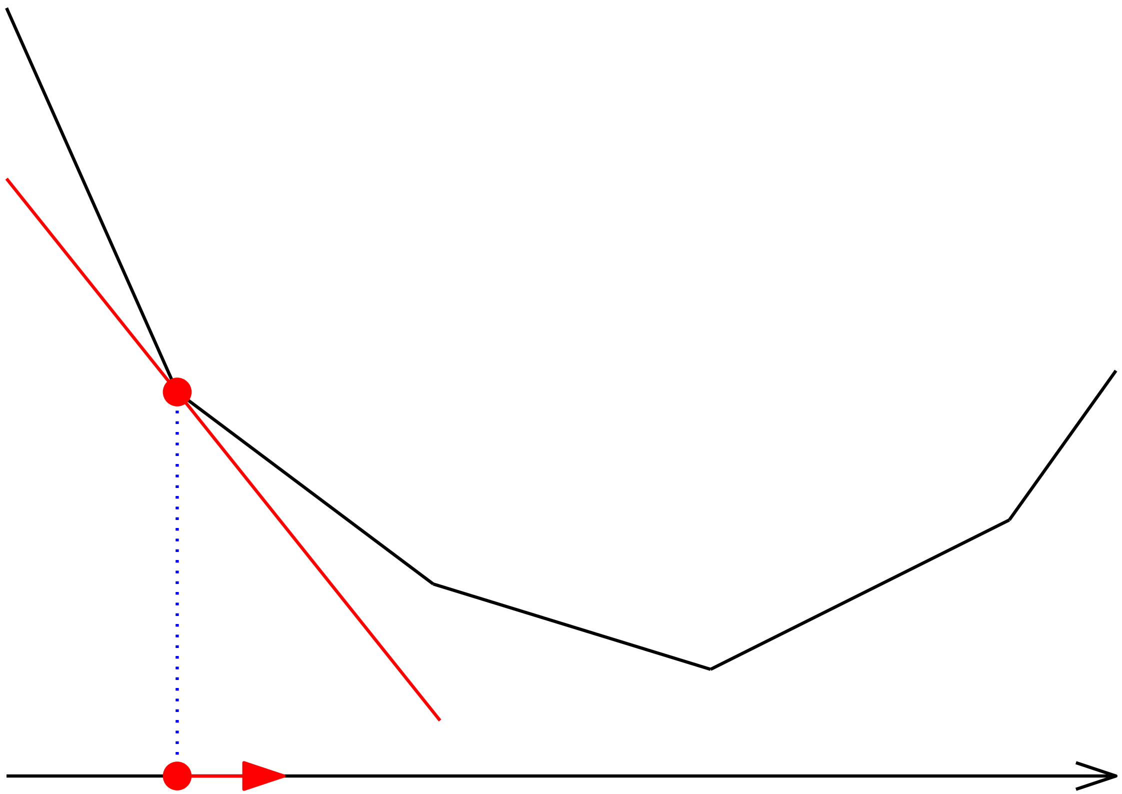
$$\nabla f(\boldsymbol{x}^{(i)}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\boldsymbol{x}^{(i)}) \\ \frac{\partial f}{\partial x_2}(\boldsymbol{x}^{(i)}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\boldsymbol{x}^{(i)}) \end{bmatrix}$$

劣勾配法は 勾配法の考え方を
微分不可能凸関数に適用したもの



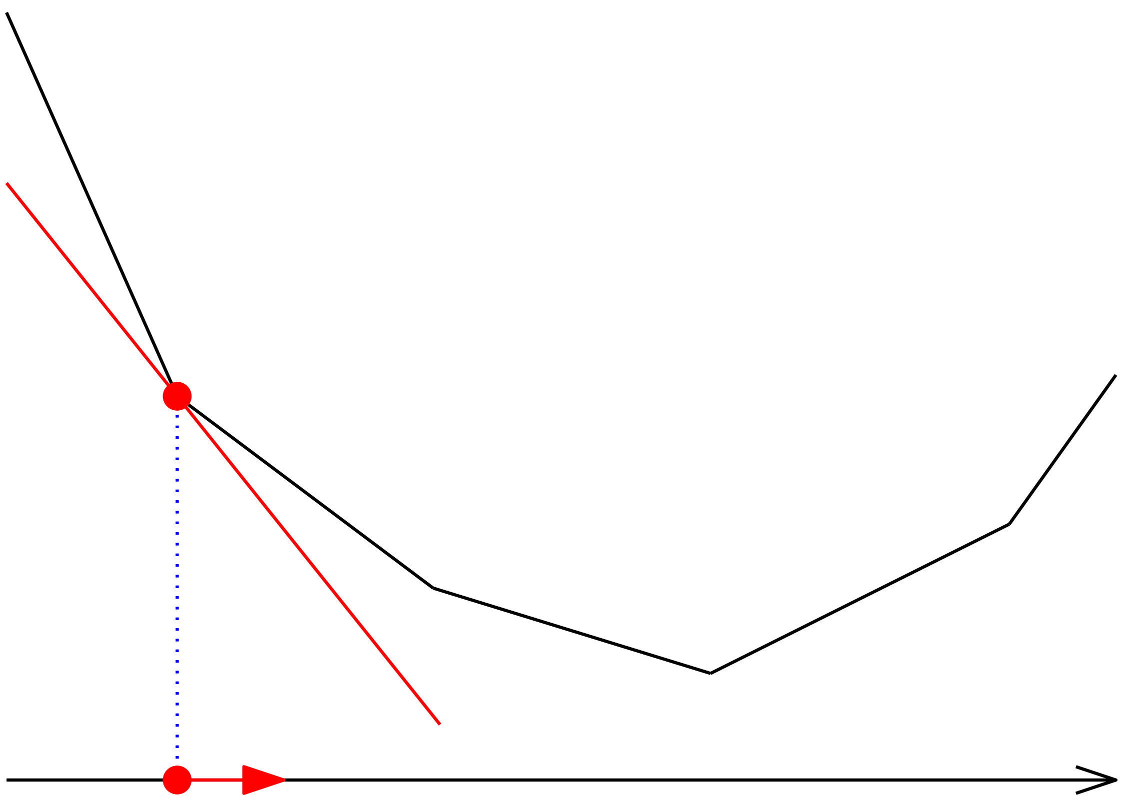
今日の内容

- ラグランジュ緩和：復習
- ラグランジュ緩和の性質
- 劣微分と劣勾配
- 劣勾配法



今日の内容

- ラグランジュ緩和：復習
- ラグランジュ緩和の性質
- 劣微分と劣勾配
- 劣勾配法



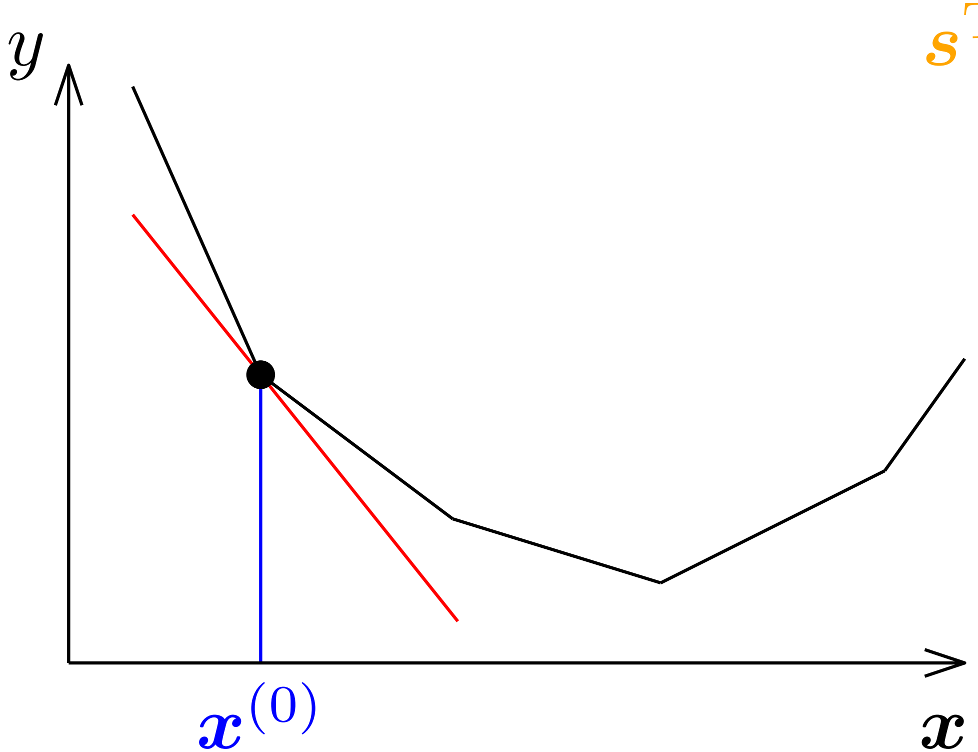
凸関数 $f\colon \mathbb{R}^n \to \mathbb{R}$, 点 $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$

定義：劣勾配 (subgradient)

ベクトル $\boldsymbol{s} \in \mathbb{R}^n$ が点 $\boldsymbol{x}^{(0)}$ における f の **劣勾配** であるとは, 次を満たすこと

$$\text{任意の}\ \boldsymbol{x} \in \mathbb{R}^n\ \text{に対して},\ \underline{\boldsymbol{s}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^{(0)}) \leq f(\boldsymbol{x})-f(\boldsymbol{x}^{(0)})}$$

$$\boldsymbol{s}^{\mathrm{T}}\boldsymbol{x} - \boldsymbol{s}^{\mathrm{T}}\boldsymbol{x}^{(0)} + f(\boldsymbol{x}^{(0)}) \leq f(\boldsymbol{x})$$



$\boxed{\text{補足}}$: f が $\boldsymbol{x}^{(0)}$ で微分可能 $\Rightarrow$

$\boldsymbol{x}^{(0)}$ における f の勾配が

$\boldsymbol{x}^{(0)}$ における f の (唯一の) 劣勾配

凸関数 $f\colon \mathbb{R}^n \to \mathbb{R}$, 点 $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$

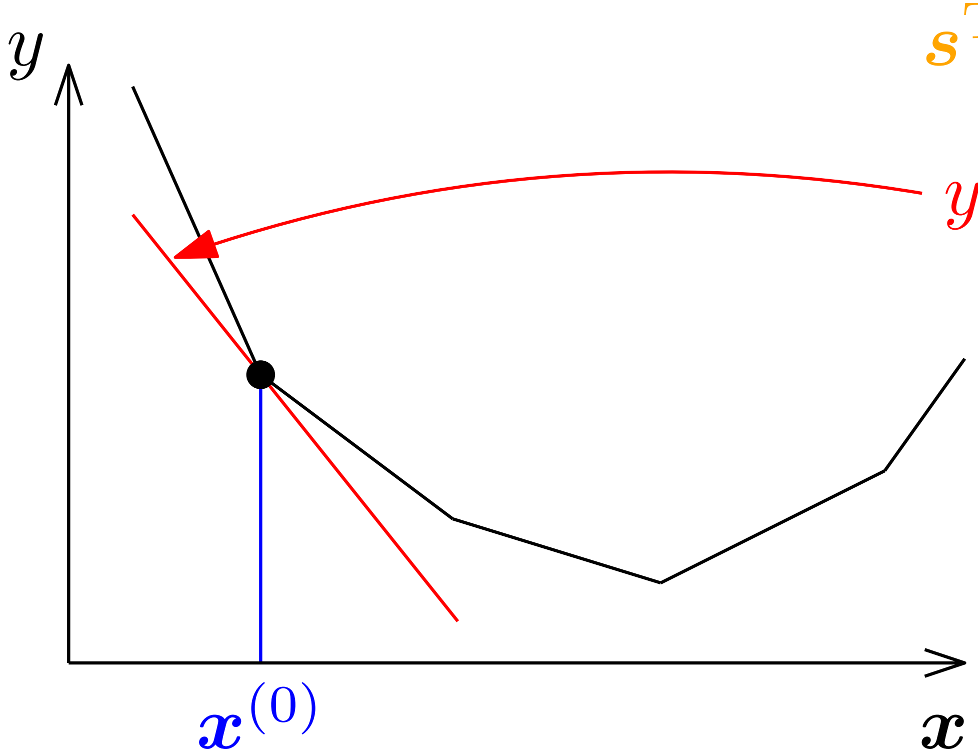
定義：劣勾配 (subgradient)

ベクトル $\boldsymbol{s} \in \mathbb{R}^n$ が点 $\boldsymbol{x}^{(0)}$ における f の **劣勾配** であるとは, 次を満たすこと

$$\text{任意の}\ \boldsymbol{x} \in \mathbb{R}^n\ \text{に対して,}\ \underline{\boldsymbol{s}^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}^{(0)}) \le f(\boldsymbol{x})-f(\boldsymbol{x}^{(0)})}$$

$$\boldsymbol{s}^{\mathrm{T}}\boldsymbol{x} - \boldsymbol{s}^{\mathrm{T}}\boldsymbol{x}^{(0)} + f(\boldsymbol{x}^{(0)}) \le f(\boldsymbol{x})$$

$$y = \boldsymbol{s}^{\mathrm{T}}\boldsymbol{x} - \boldsymbol{s}^{\mathrm{T}}\boldsymbol{x}^{(0)} + f(\boldsymbol{x}^{(0)})$$



補足 : f が $\boldsymbol{x}^{(0)}$ で微分可能 $\Rightarrow$

$\boldsymbol{x}^{(0)}$ における f の勾配が

$\boldsymbol{x}^{(0)}$ における f の (唯一の) 劣勾配

劣微分

凸関数 $f\colon \mathbb{R}^n \to \mathbb{R}$，点 $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$

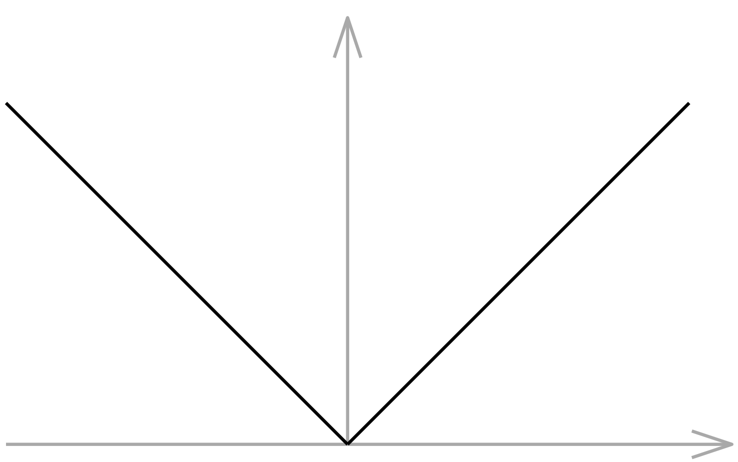
定義：劣微分 (subdifferential)

点 $\boldsymbol{x}^{(0)}$ における f の **劣微分** とは，
$\boldsymbol{x}^{(0)}$ における f の劣勾配をすべて集めた集合のこと

$$\underline{\partial f(\boldsymbol{x}^{(0)})} = \{\boldsymbol{s} \in \mathbb{R}^n \mid \boldsymbol{s} \text{ は } \boldsymbol{x}^{(0)} \text{ における } f \text{ の劣勾配}\}$$

劣微分の記法

例：$n = 1$，$f(x) = |x|$ のとき



$$\partial f(x) = \begin{cases} \{-1\} & (x < 0 \text{ のとき}) \\ [-1, 1] & (x = 0 \text{ のとき}) \\ \{1\} & (x > 0 \text{ のとき}) \end{cases}$$

劣勾配，劣微分と最小解

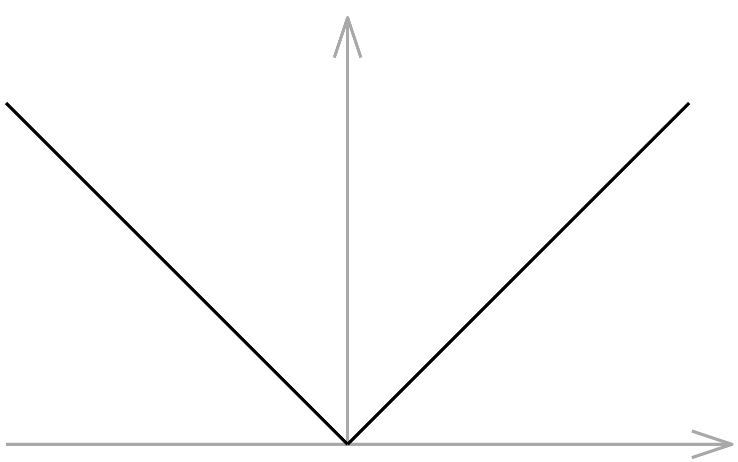
凸関数 $f\colon \mathbb{R}^n \to \mathbb{R}$，点 $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$

性質：劣微分と最小解

$\boldsymbol{x}^{(0)}$ が f の最小解 $\quad\Leftrightarrow\quad \mathbf{0} \in \partial f(\boldsymbol{x}^{(0)})$

微分可能である場合と同様に証明できる (微分積分学の内容を参照)

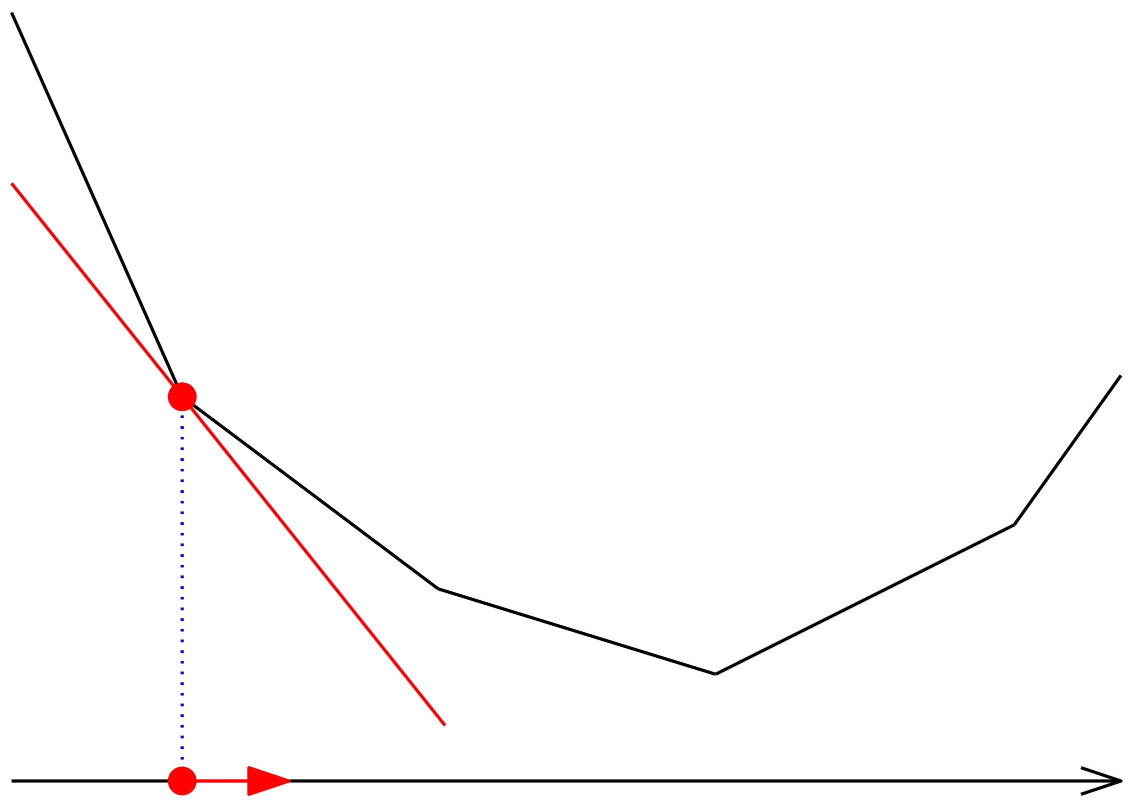
例：$n=1$，$f(x) = |x|$ のとき



$$\partial f(x) = \begin{cases} \{-1\} & (x < 0 \text{ のとき}) \\ [-1, 1] & (x = 0 \text{ のとき}) \\ \{1\} & (x > 0 \text{ のとき}) \end{cases}$$

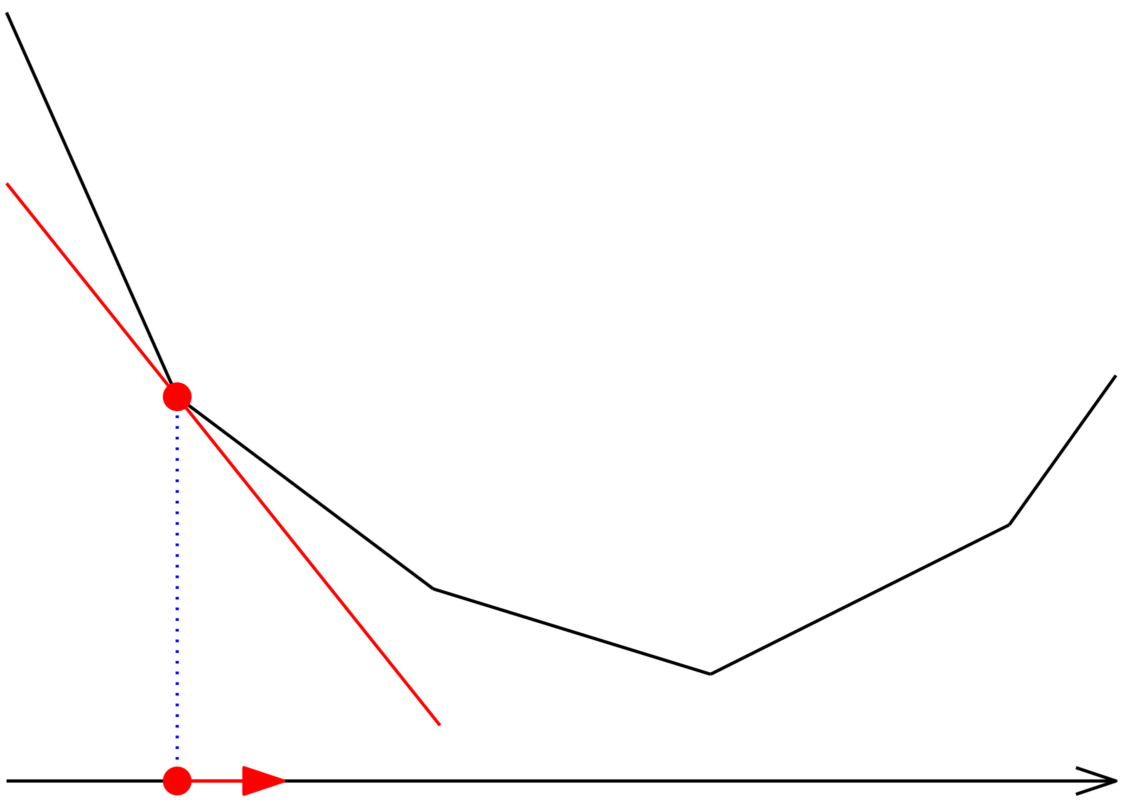
今日の内容

- ラグランジュ緩和：復習
- ラグランジュ緩和の性質
- 劣微分と劣勾配
- 劣勾配法



今日の内容

- ラグランジュ緩和：復習
- ラグランジュ緩和の性質
- 劣微分と劣勾配
- $\boxed{\text{劣勾配法}}$



劣勾配法

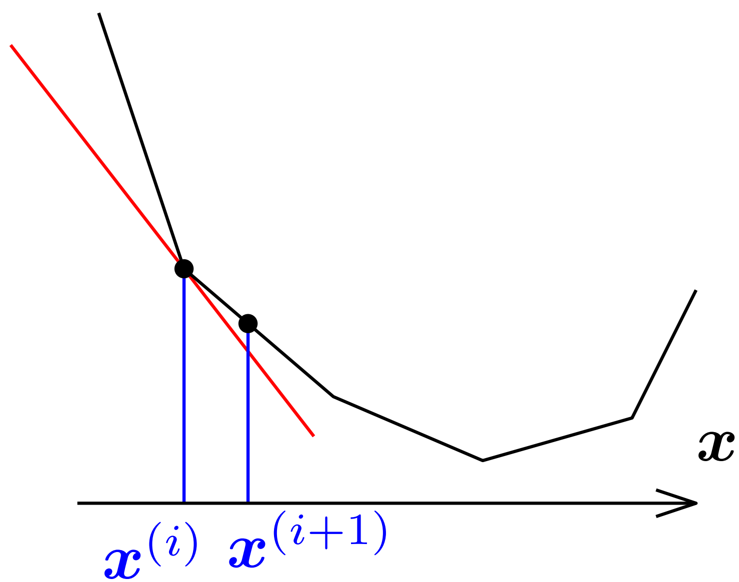
凸関数 $f\colon \mathbb{R}^n \to \mathbb{R}$ (微分可能であるとは限らない)

アルゴリズム：劣勾配法 (subgradient method)

1. 初期点 $\boldsymbol{x}^{(0)}$ を適当に選ぶ
2. $i = 0, 1, \ldots$ で収束するまで以下を繰り返す

$$\boldsymbol{s}^{(i)} \in \partial f(\boldsymbol{x}^{(i)}),$$
$$\boldsymbol{x}^{(i+1)} := \boldsymbol{x}^{(i)} - \alpha_i \boldsymbol{s}^{(i)}$$

($\alpha_i > 0$ はうまく選んだ実数)



課題

1. 劣勾配 $\boldsymbol{s}^{(i)}$ の見つけ方
2. 係数 α_i の定め方

ステップサイズ とよく呼ばれる

係数 α_i の決め方

いろいろな決め方があるが，理論的には次が古くから知られている

性質：ステップサイズの選び方 (Polyak '67)

α_i は次を満たす数列であるとする

- $\displaystyle\lim_{i\to\infty} \alpha_i = 0, \sum_{i=0}^{\infty} \alpha_i = \infty$

$\boldsymbol{s}^{(i)}$ は次を満たすとする

- 数列 $\{\|\boldsymbol{s}^{(i)}\|^2\}$ が有界

このとき，$\boldsymbol{x}^{(0)}, \boldsymbol{x}^{(1)}, \ldots$ は f の最小解に収束する

ただ，このステップサイズでは収束が遅い ことが知られている

実用上は，経験的に収束が速いステップサイズを用いることも多い
(例：指数的に減少するステップサイズ $\alpha_{i+1} = r\alpha_i$)

劣勾配 $\boldsymbol{s}^{(i)}$ の決め方

ラグランジュ緩和の最適値関数 g の場合

$$-g(\boldsymbol{\lambda}) = -\min\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \mid A_1\boldsymbol{x} = \boldsymbol{b}_1, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathbb{Z}^n\}$$

最適解を $x_{\boldsymbol{\lambda}}$ とする

性質：最適値関数 $-g$ の劣勾配

$$A_2\boldsymbol{x}_{\boldsymbol{\lambda}} - \boldsymbol{b}_2 \in \partial(-g)(\boldsymbol{\lambda})$$

証明：次を証明する

任意の $\boldsymbol{\mu} \in \mathbb{R}^m$ に対して，$(A_2\boldsymbol{x}_{\boldsymbol{\lambda}} - \boldsymbol{b}_2)^{\mathrm{T}}(\boldsymbol{\mu} - \boldsymbol{\lambda}) \leq -g(\boldsymbol{\mu}) + g(\boldsymbol{\lambda})$

劣勾配 $\boldsymbol{s}^{(i)}$ の決め方

ラグランジュ緩和の最適値関数 g の場合

$$-g(\boldsymbol{\lambda}) = -\min\{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} + \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}) \mid A_1\boldsymbol{x} = \boldsymbol{b}_1, \boldsymbol{x} \geq \boldsymbol{0}, \boldsymbol{x} \in \mathbb{Z}^n\}$$

→ 最適解を $\boldsymbol{x}_{\boldsymbol{\lambda}}$ とする

性質：最適値関数 $-g$ の劣勾配

$$A_2\boldsymbol{x}_{\boldsymbol{\lambda}} - \boldsymbol{b}_2 \in \partial(-g)(\boldsymbol{\lambda})$$

証明：次を証明する

任意の $\boldsymbol{\mu} \in \mathbb{R}^m$ に対して，$(A_2\boldsymbol{x}_{\boldsymbol{\lambda}} - \boldsymbol{b}_2)^{\mathrm{T}}(\boldsymbol{\mu} - \boldsymbol{\lambda}) \leq -g(\boldsymbol{\mu}) + g(\boldsymbol{\lambda})$

$$\begin{aligned}
&-g(\boldsymbol{\lambda}) + (A_2\boldsymbol{x}_{\boldsymbol{\lambda}} - \boldsymbol{b}_2)^{\mathrm{T}}(\boldsymbol{\mu} - \boldsymbol{\lambda}) \\
&\quad = -\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\lambda}} - \boldsymbol{\lambda}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\lambda}}) + (A_2\boldsymbol{x}_{\boldsymbol{\lambda}} - \boldsymbol{b}_2)^{\mathrm{T}}(\boldsymbol{\mu} - \boldsymbol{\lambda}) \\
&\quad = -\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{\boldsymbol{\lambda}} - \boldsymbol{\mu}^{\mathrm{T}}(\boldsymbol{b}_2 - A_2\boldsymbol{x}_{\boldsymbol{\lambda}}) \\
&\quad \leq -g(\boldsymbol{\mu})
\end{aligned}$$

□

微分不可能凸関数の最小化：補足

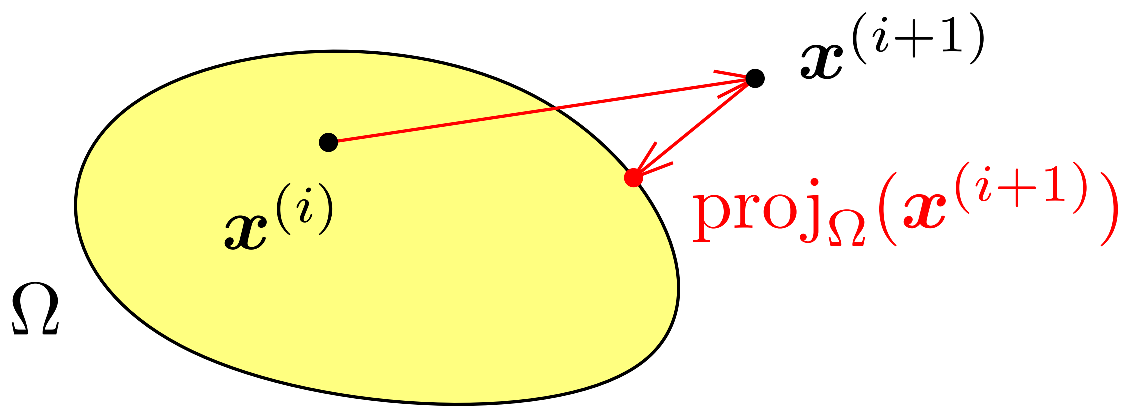
劣勾配法以外のアルゴリズム

- 平滑化法 (Nesterov ’05)
- FISTA (Beck, Teboulle ’09)
- $\cdots$

始域が $\mathbb{R}^n$ ではなく，凸集合 Ω の場合

計算した $\boldsymbol{x}^{(i+1)}$ が Ω をはみ出ることがある

- Ω に引き戻す必要がある (射影)



次回予告

次回の内容

前求解

- 問題のサイズを小さくする／問題を解きやすくする

授業評価アンケートに回答する時間も設ける