連続最適化(2):
制約付き最適化

佐藤　一誠
sato@k.u-tokyo.ac.jp
http://www.ms.k.u-tokyo.ac.jp

制約付き最適化問題

- 制約付き最適化問題(constrained optimization problem): $\mathcal{X} \subset \mathbb{R}^d$ 上に定義される関数 $f : \mathcal{X} \to \mathbb{R}$ の最小値を求めよ

$$\min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0},\ \boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))^\top$$

$$\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), \ldots, h_n(\boldsymbol{x}))^\top$$

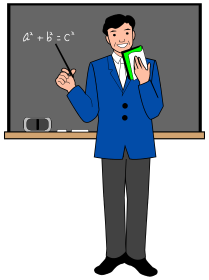
- “subject to…”は「…を条件として」という意味
- ベクトルの等式・不等式は要素毎に適用
- g_i, h_i は互いに一次独立(linearly independent)であると仮定

実行可能性と凸性

$$\min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x})$$

$$\text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0},\ \boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0}$$

- $\boldsymbol{x}$ が全ての制約条件を満たすとき，$\boldsymbol{x}$ を実行可能解(feasible solution)という
  - 最適解は，実行可能解の中で目的関数を最小にする
- 凸最適化問題(convex optimization problem)：関数 f が凸，集合 $\mathcal{X}$ が凸，制約条件 $\boldsymbol{g}, \boldsymbol{h}$ が凸
  - 最適値が一意に定まる
- 以下，凸最適化問題を考える



講義の流れ

1. 等式制約付き最適化問題
   - A) ラグランジュ未定乗数法
   - B) 双対上昇法
   - C) 乗数法
2. 不等式制約付き最適化問題

等式制約付き最適化問題

等式制約付き最適化問題

$$\min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))^\top$$

罰則法(penalty method):

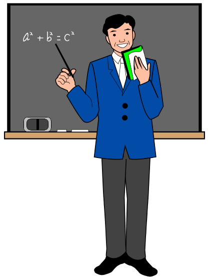
- 初期値を $\boldsymbol{x}_k$ に設定して，勾配法や準ニュートン法などで以下の制約なし最適化問題を解く:

$$\boldsymbol{x}_{k+1} = \operatorname*{argmin}_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) + c_k \|\boldsymbol{g}(\boldsymbol{x})\|^2$$

- c_k の値を徐々に増やしていく:

$$0 < c_1 < c_2 < \cdots$$

c_k の決め方が難しい



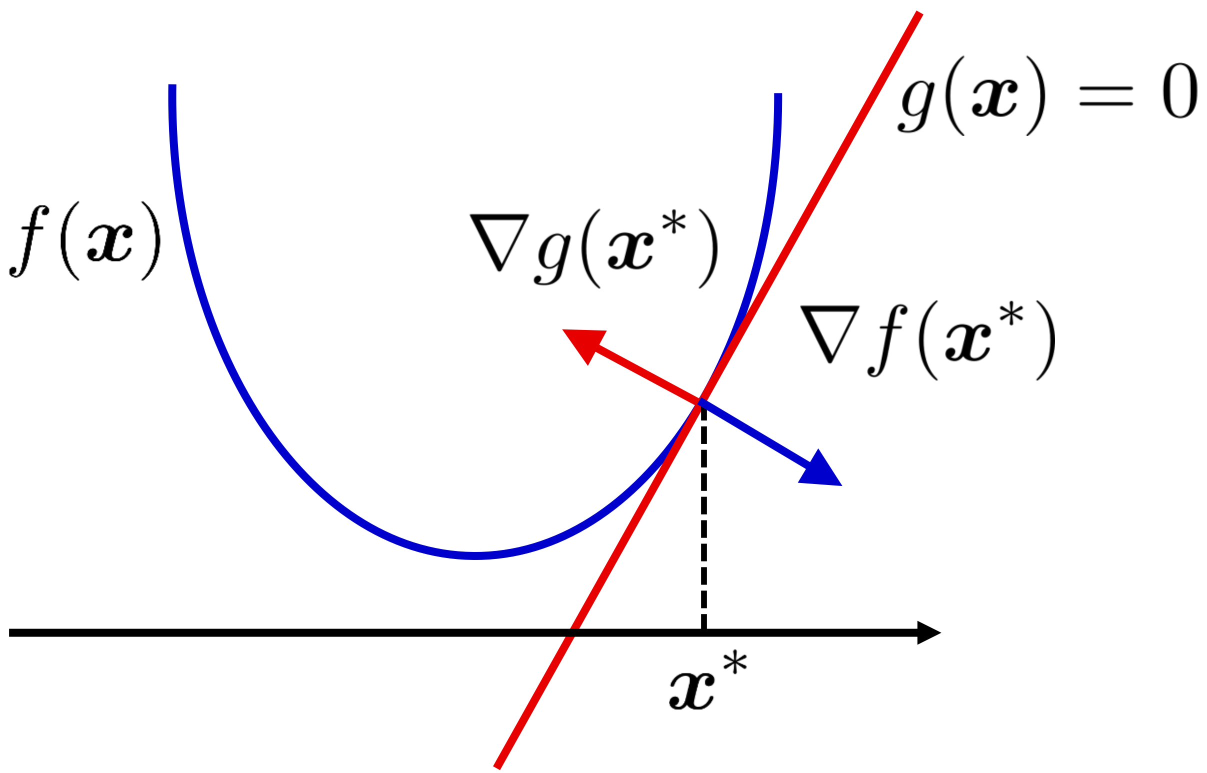
講義の流れ

1. 等式制約付き最適化問題
   - A) ラグランジュ未定乗数法
   - B) 双対上昇法
   - C) 乗数法
   - D) 交互方向乗数法
2. 不等式制約付き最適化問題

最適性条件

$$\min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } g(\boldsymbol{x}) = 0$$

- 最適解 $\boldsymbol{x}^*$では，$\nabla f(\boldsymbol{x}^*)$と $\nabla g(\boldsymbol{x}^*)$ が平行になる



最適性条件

$$\min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))^\top$$

- 凸最適化問題に対して，$\boldsymbol{x}^*$ が最適解となるための必要条件は，ある $\boldsymbol{\lambda}^*$ に対して，以下の2条件が成り立つこと:
  - $\nabla f(\boldsymbol{x}^*) + \nabla \boldsymbol{g}(\boldsymbol{x}^*)^\top \boldsymbol{\lambda}^* = \boldsymbol{0}$

$$\nabla \boldsymbol{g}(\boldsymbol{x}) = (\nabla g_1(\boldsymbol{x}), \ldots, \nabla g_m(\boldsymbol{x}))^\top$$

m個の等式制約

  - $\boldsymbol{g}(\boldsymbol{x}^*) = \boldsymbol{0}$

$$\nabla f(\boldsymbol{x}^*) + \sum_{j=1}^{m} \lambda_j^* \nabla g_j(\boldsymbol{x}^*) = \boldsymbol{0}$$

ラグランジュ未定乗数法

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$$

- ラグランジュ関数(Lagrangian):

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x})$$

- 最適解は以下の方程式を満たす(最適性条件):

$$\nabla_{\boldsymbol{\lambda}} L(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{0}$$

$$\nabla_{\boldsymbol{x}} L(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{0}$$

- この方程式を解くことによって解を求める方法をラグランジュ未定乗数法(method of Lagrange multipliers)とよぶ

演習

- 以下の最適化問題の解を，ラグランジュ未定乗数法を用いて求めよ

$$\min_{x,y\in\mathbb{R}} 3x^2 + 2y^2$$

$$\text{subject to } x + y = 1$$

解答例

$$\min_{x,y\in\mathbb{R}} \ 3x^2 + 2y^2 \quad \text{subject to } x + y = 1$$

- ラグランジュ関数は

$$L(x, y, \lambda) = 3x^2 + 2y^2 + \lambda(x + y - 1)$$

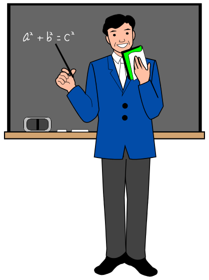
- これより,

$$\frac{\partial L(x, y, \lambda)}{\partial x} = 6x + \lambda = 0$$

$$\frac{\partial L(x, y, \lambda)}{\partial y} = 4y + \lambda = 0$$

$$\frac{\partial L(x, y, \lambda)}{\partial \lambda} = x + y - 1 = 0$$

- これらを解けば $(x, y) = (0.4, 0.6)$



講義の流れ

1. 等式制約付き最適化問題
   - A) ラグランジュ未定乗数法
   - B) 双対上昇法
   - C) 乗数法
2. 不等式制約付き最適化問題

ラグランジュ双対問題

$$f* = \min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \mathbf{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))^\top$$

■ラグランジュ双対問題(Lagrange dual problem):

$$f^* = \max_{\boldsymbol{\lambda} \in \mathbb{R}^m} \omega(\boldsymbol{\lambda}) \qquad \omega(\boldsymbol{\lambda}) = \inf_{\boldsymbol{x} \in \mathcal{X}} L(\boldsymbol{x}, \boldsymbol{\lambda})$$

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x})$$

- 凸最適化問題に対しては最適解が一致
- 双対問題は制約なしになる

演習

- 以下の最適化問題のラグランジュ双対問題を求め，最適解を求めよ

$$\min_{x,y\in\mathbb{R}} \; 3x^2 + 2y^2$$

$$\text{subject to } x + y = 1$$

解答例

$$\min_{x,y\in\mathbb{R}} 3x^2 + 2y^2 \quad \text{subject to } x + y = 1$$

- $L(x, y, \lambda) = 3x^2 + 2y^2 + \lambda(x + y - 1)$
  - $\dfrac{\partial L(x, y, \lambda)}{\partial x} = 6x + \lambda = 0 \;\Rightarrow\; x = -\dfrac{\lambda}{6}$
  - $\dfrac{\partial L(x, y, \lambda)}{\partial y} = 4y + \lambda = 0 \;\Rightarrow\; y = -\dfrac{\lambda}{4}$
- $\omega(\lambda) = \inf_{x,y\in\mathbb{R}} L(x, y, \lambda) = -\dfrac{5}{24}\lambda^2 - \lambda$
- $\operatorname{argmax}_{\lambda\in\mathbb{R}} \omega(\lambda) = -\dfrac{12}{5} \;\Rightarrow\; (x, y) = (0.4, 0.6)$

双対変数の最適化

$$f^* = \max_{\boldsymbol{\lambda}\in\mathbb{R}^m} \omega(\boldsymbol{\lambda}) \qquad \omega(\boldsymbol{\lambda}) = \inf_{\boldsymbol{x}\in\mathcal{X}} L(\boldsymbol{x},\boldsymbol{\lambda})$$

$$L(\boldsymbol{x},\boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x})$$

- 双対変数 $\boldsymbol{\lambda}$ を勾配法で最適化:

$$\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \varepsilon_k \nabla\omega(\boldsymbol{\lambda}_k)$$

$\varepsilon_k > 0$:ステップ幅

- $\boldsymbol{x}_{k+1} = \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} L(\boldsymbol{x},\boldsymbol{\lambda}_k)$ とおけば,

$\omega(\boldsymbol{\lambda}_k) = f(\boldsymbol{x}_{k+1}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x}_{k+1})$ より

$$\nabla\omega(\boldsymbol{\lambda}_k) = \boldsymbol{g}(\boldsymbol{x}_{k+1})$$

双対上昇法

$$f^* = \max_{\boldsymbol{\lambda} \in \mathbb{R}^m} \omega(\boldsymbol{\lambda}) \qquad \omega(\boldsymbol{\lambda}) = \inf_{\boldsymbol{x} \in \mathcal{X}} L(\boldsymbol{x}, \boldsymbol{\lambda})$$

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x})$$

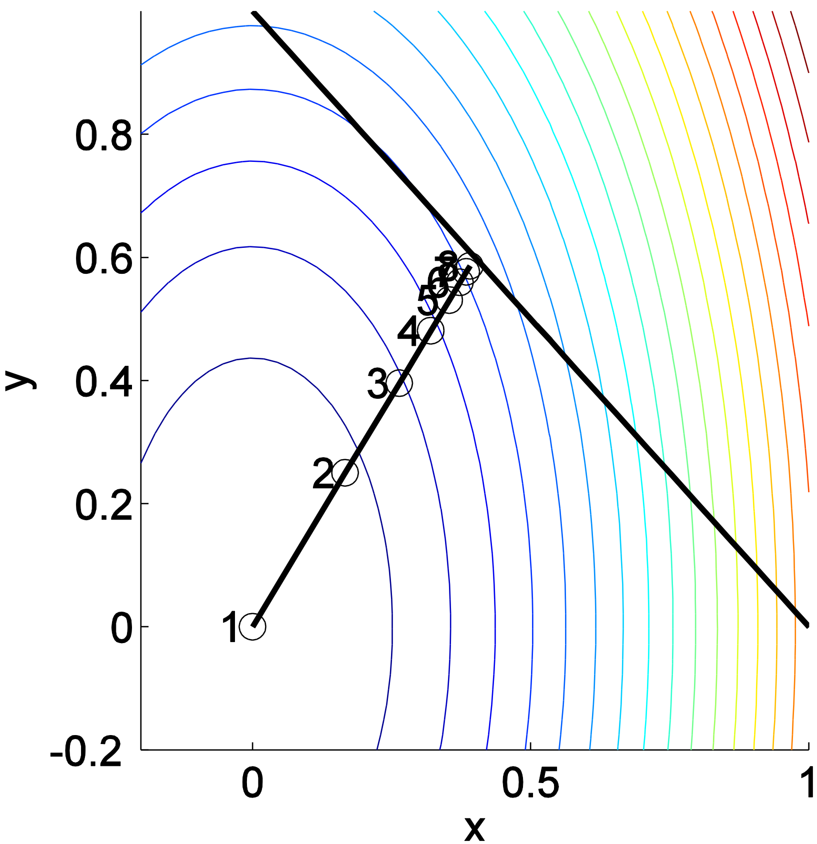
- 双対上昇法(dual ascent method) : 適当な初期値から, ラグランジュ関数の最小化と双対変数の勾配上昇を交互に実行
  - $\boldsymbol{x}_{k+1} = \operatorname*{argmin}_{\boldsymbol{x} \in \mathcal{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}_k)$
  - $\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \varepsilon_k \boldsymbol{g}(\boldsymbol{x}_{k+1})$

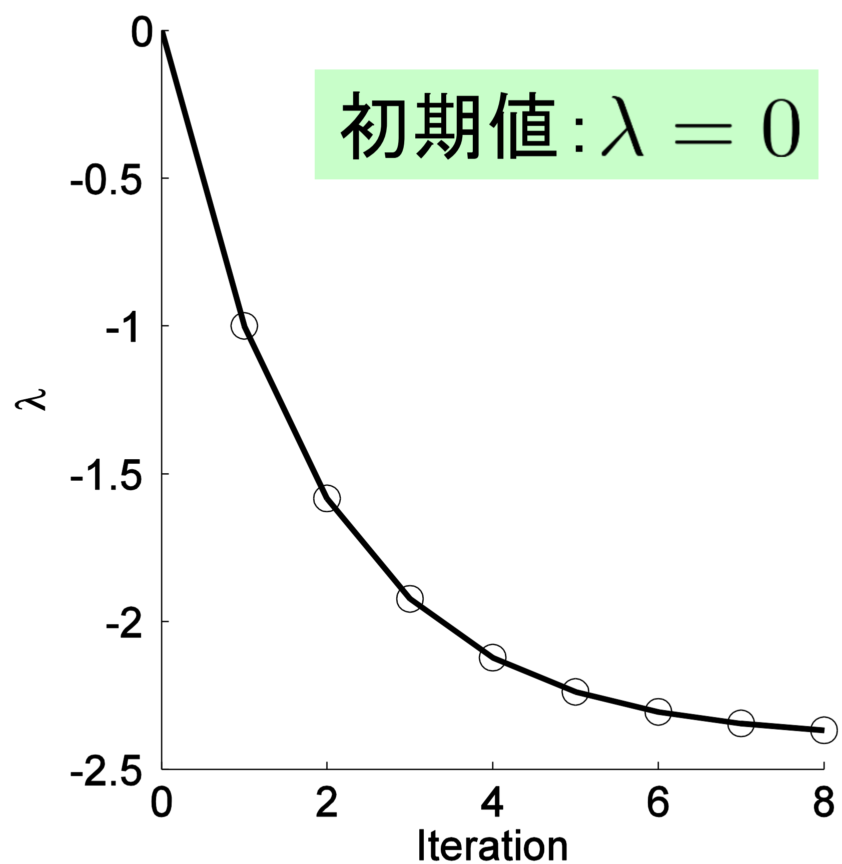
$\varepsilon_k > 0$: ステップ幅

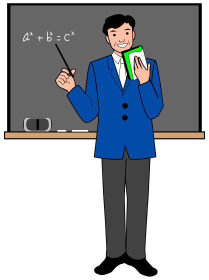
実行例

$$\min_{x,y\in\mathbb{R}} 3x^2 + 2y^2 \quad \text{subject to } x + y = 1$$

- 最適解 $(x, y) = (0.4, 0.6)$, $\lambda = -\frac{12}{5}$ に収束







講義の流れ

1. 等式制約付き最適化問題
   - A) ラグランジュ未定乗数法
   - B) 双対上昇法
   - C) 乗数法
2. 不等式制約付き最適化問題

拡張ラグランジュ関数

- 拡張ラグランジュ関数(augmented Lagrangian):

$$L_c(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x}) + \frac{c}{2}\|\boldsymbol{g}(\boldsymbol{x})\|^2 \quad c \geq 0$$

  - 二次の項を加える事により凸性が増す
- 双対上昇法と同様に，双対変数 $\boldsymbol{\lambda}$ を勾配法で最適化：

$$\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \varepsilon_k \nabla \omega_c(\boldsymbol{\lambda}_k)$$

$\varepsilon_k > 0$:ステップ幅

$$\omega_c(\boldsymbol{\lambda}) = \inf_{\boldsymbol{x} \in \mathcal{X}} L_c(\boldsymbol{x}, \boldsymbol{\lambda})$$

双対変数の最適化

$$\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \varepsilon_k \nabla \omega_c(\boldsymbol{\lambda}_k)$$

$$L_c(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x}) + \frac{c}{2}\|\boldsymbol{g}(\boldsymbol{x})\|^2 \qquad c \geq 0$$

- $\boldsymbol{x}_{k+1} = \underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmin}}\, L_c(\boldsymbol{x}, \boldsymbol{\lambda}_k)$ とおけば,

$$\begin{aligned}\omega_c(\boldsymbol{\lambda}_k) &= \inf_{\boldsymbol{x}\in\mathcal{X}} L_c(\boldsymbol{x}, \boldsymbol{\lambda}_k) \\ &= f(\boldsymbol{x}_{k+1}) + \boldsymbol{\lambda}_k^\top \boldsymbol{g}(\boldsymbol{x}_{k+1}) + \frac{c}{2}\|\boldsymbol{g}(\boldsymbol{x}_{k+1})\|^2\end{aligned}$$

より $\nabla \omega_c(\boldsymbol{\lambda}_k) = \boldsymbol{g}(\boldsymbol{x}_{k+1})$

乗数法

- 乗数法(method of multipliers):
  適当な初期値から,
  拡張ラグランジュ関数の最小化と
  双対変数の勾配上昇を交互に実行:
  - $\boldsymbol{x}_{k+1} = \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} L_c(\boldsymbol{x}, \boldsymbol{\lambda}_k)$
  - $\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + c\boldsymbol{g}(\boldsymbol{x}_{k+1})$
- ステップ幅は $\varepsilon_k = c$ と設定すると,

$$\boldsymbol{x}_{k+1} = \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}_{k+1})$$

の最適性条件が満たされる(次ページ参照)

乗数法のステップ幅の選択

$$L_c(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x}) + \frac{c}{2}\|\boldsymbol{g}(\boldsymbol{x})\|^2$$

$$\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + c\boldsymbol{g}(\boldsymbol{x}_{k+1})$$

- $\boldsymbol{x}_{k+1} = \underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmin}} L_c(\boldsymbol{x}, \boldsymbol{\lambda}_k)$ の最適性条件より

$$\begin{aligned}
\boldsymbol{0} &= \nabla_{\boldsymbol{x}} L_c(\boldsymbol{x}_{k+1}, \boldsymbol{\lambda}_k) \\
&= \nabla_{\boldsymbol{x}} f(\boldsymbol{x}_{k+1}) + \nabla_{\boldsymbol{x}} \boldsymbol{g}(\boldsymbol{x}_{k+1})^\top (\boldsymbol{\lambda}_k + c\boldsymbol{g}(\boldsymbol{x}_{k+1})) \\
&= \nabla_{\boldsymbol{x}} f(\boldsymbol{x}_{k+1}) + \nabla_{\boldsymbol{x}} \boldsymbol{g}(\boldsymbol{x}_{k+1})^\top \boldsymbol{\lambda}_{k+1} \\
&= \nabla_{\boldsymbol{x}} L(\boldsymbol{x}_{k+1}, \boldsymbol{\lambda}_{k+1})
\end{aligned}$$

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x})$$

➡ $\boldsymbol{x}_{k+1} = \underset{\boldsymbol{x}\in\mathcal{X}}{\operatorname{argmin}} L(\boldsymbol{x}, \boldsymbol{\lambda}_{k+1})$ の最適性条件も満たす

実行例

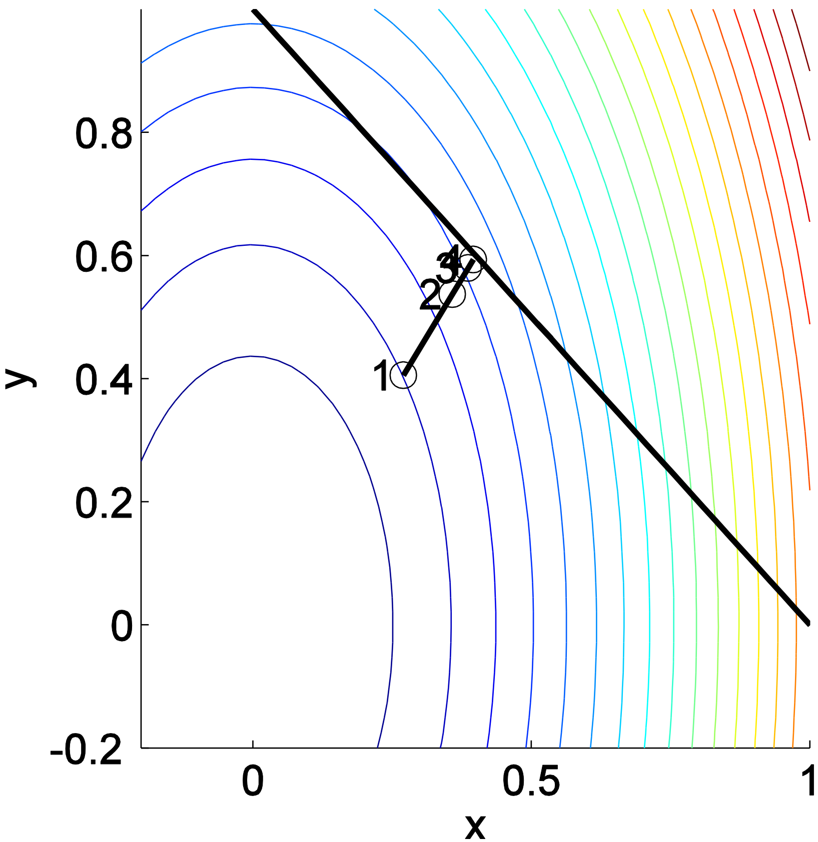
$$\min_{x,y\in\mathbb{R}} 3x^2 + 2y^2 \quad \text{subject to } x + y = 1$$

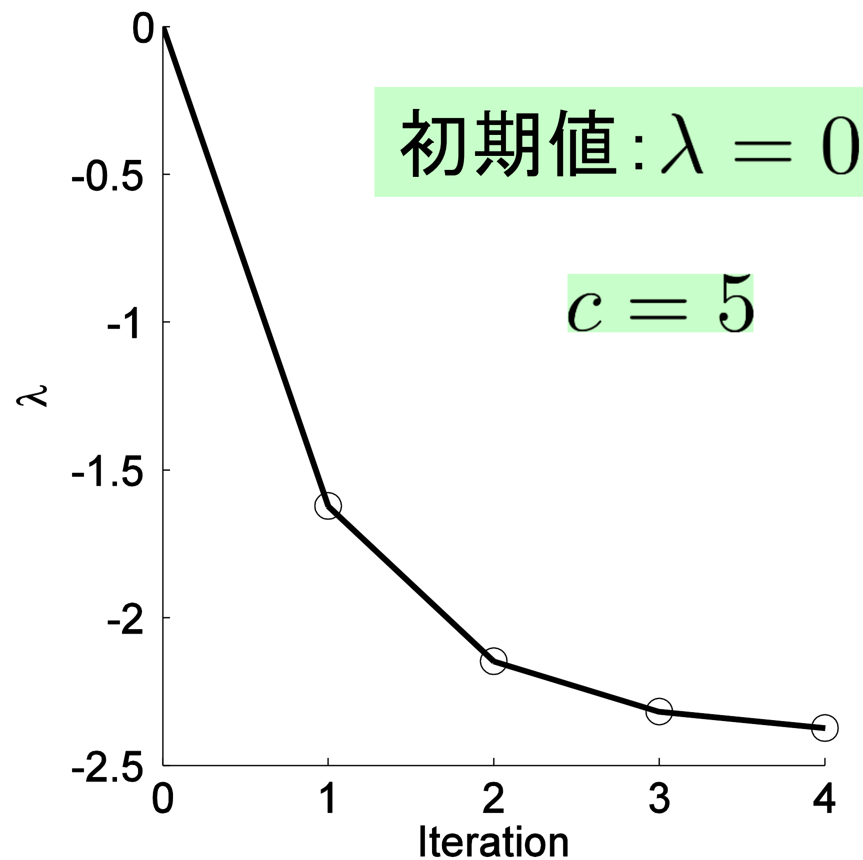
- $L_c(x,y,\lambda) = 3x^2 + 2y^2 + \lambda(x+y-1) + \frac{c}{2}(x+y-1)^2$
  - $\frac{\partial L_c(x,y,\lambda)}{\partial x} = 6x + \lambda + c(x+y-1) = 0$ ⇒ $x = \frac{c-\lambda}{6+5c/2}$
  - $\frac{\partial L_c(x,y,\lambda)}{\partial y} = 4y + \lambda + c(x+y-1) = 0$ ⇒ $y = \frac{c-\lambda}{4+5c/3}$
- 乗数法のアルゴリズム:
  - $x_{k+1} = \frac{c-\lambda_k}{6+5c/2}$, $y_{k+1} = \frac{c-\lambda_k}{4+5c/3}$
  - $\lambda_{k+1} = \lambda_k + c(x+y-1)$

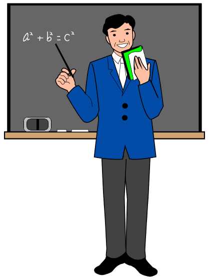
実行例

$$\min_{x,y\in\mathbb{R}} 3x^2 + 2y^2 \quad \text{subject to } x + y = 1$$

■この例では双対上昇法より速く収束







講義の流れ

1. 等式制約付き最適化問題
2. 不等式制約付き最適化問題
   A) 最適性条件
   B) 双対問題

不等式制約付き最適化問題

- 不等式制約付き最適化問題

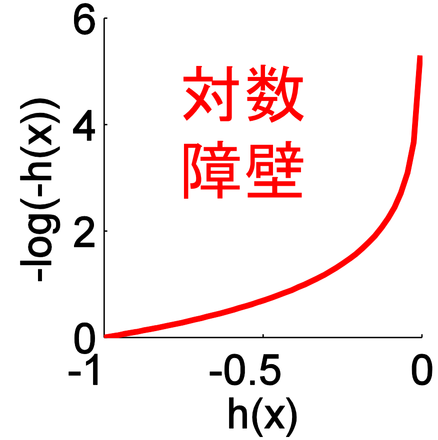
$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0}$$

$$\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), \ldots, h_n(\boldsymbol{x}))^\top$$

- $h_i(\boldsymbol{x}) = 0$ のとき，この制約を有効制約 (active constraint)とよぶ

制約なし最適化問題への変換

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } h(\boldsymbol{x}) \leq 0$$



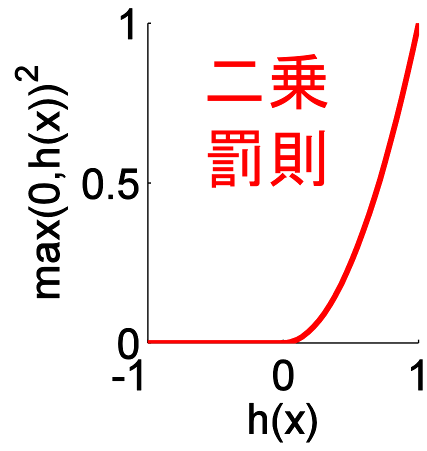
- 障壁法(barrier method):
  - 初期値を $\boldsymbol{x}_k$ に設定して，勾配法や準ニュートン法などで以下の制約なし最適化問題を解く:

$$\boldsymbol{x}_{k+1} = \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) - c_k \log(-h(\boldsymbol{x}))$$

  - c_k の値を徐々に減らしていく: $c_1 > c_2 > \cdots > 0$
- 途中の解 $\boldsymbol{x}_{k+1}$ が常に $h(\boldsymbol{x}_{k+1}) \leq 0$ を満たす（つまり，実行可能領域の内側にある）ことから，内点法(interior-point method)ともよばれる
- c_k の決め方が難しい

制約なし最適化問題への変換

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } h(\boldsymbol{x}) \leq 0$$



- 罰則法(penalty method):
  - 初期値を$\boldsymbol{x}_k$に設定して,
    勾配法や準ニュートン法などで
    以下の制約なし最適化問題を解く:

$$\boldsymbol{x}_{k+1} = \operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) + c_k \max(0, h(\boldsymbol{x}))^2$$

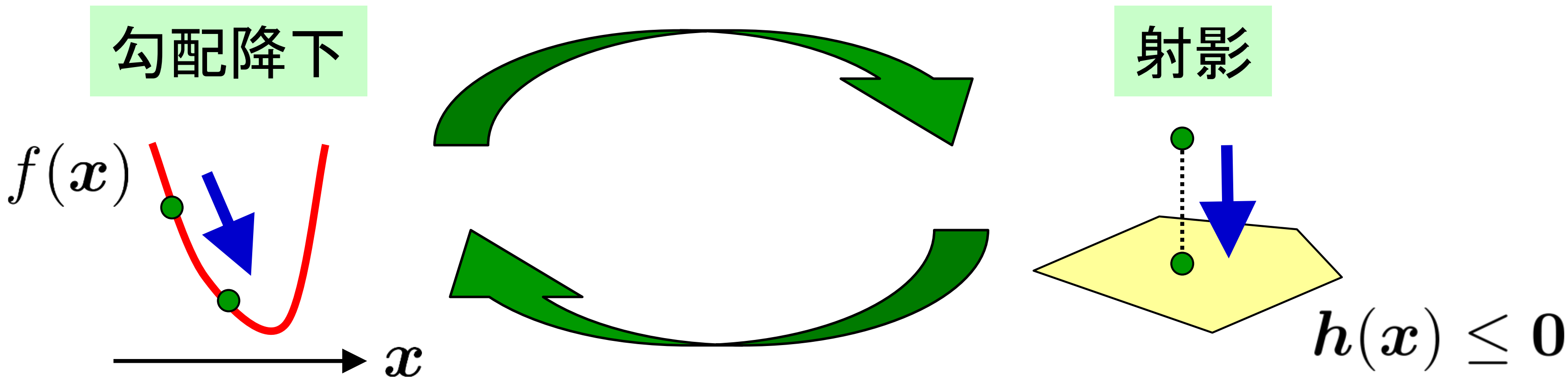
  - c_kの値を徐々に増やしていく: $0 < c_1 < c_2 < \cdots$
- 途中の解 $\boldsymbol{x}_{k+1}$ が一般に $h(\boldsymbol{x}_{k+1}) > 0$ を満たす(つまり, 実行可能領域の外側にある)ことから, 外点法(exterior-point method)ともよばれる
- c_kの決め方が難しい

射影勾配法

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0}$$

■ 射影勾配法(projected gradient method):

- 勾配降下と実行可能領域への射影を繰り返す

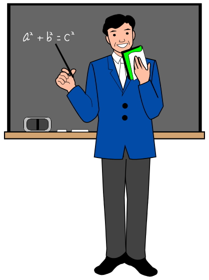


$$\widetilde{\boldsymbol{x}}_{k+1} = \boldsymbol{x}_k - \varepsilon_k \nabla f(\boldsymbol{x}_k)$$

$$\varepsilon_k > 0$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{P}_k \widetilde{\boldsymbol{x}}_{k+1}$$

- まとめると $\boldsymbol{x}_{k+1} = \boldsymbol{P}_k(\boldsymbol{x}_k - \varepsilon_k \nabla f(\boldsymbol{x}_k))$
- 射影が簡単に計算できるとき，効率が良い



講義の流れ

1. 等式制約付き最適化問題
2. 不等式制約付き最適化問題
   A) 最適性条件
   B) 双対問題

有効な制約と有効でない制約

$$\min_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } h(\boldsymbol{x}) \leq 0$$

- 許容解xにおいて$h(x)=0$である制約を有効な制約(active constraint)

- 許容解xにおいて$h(x)<0$である制約を有効でない制約(inactive constraint)

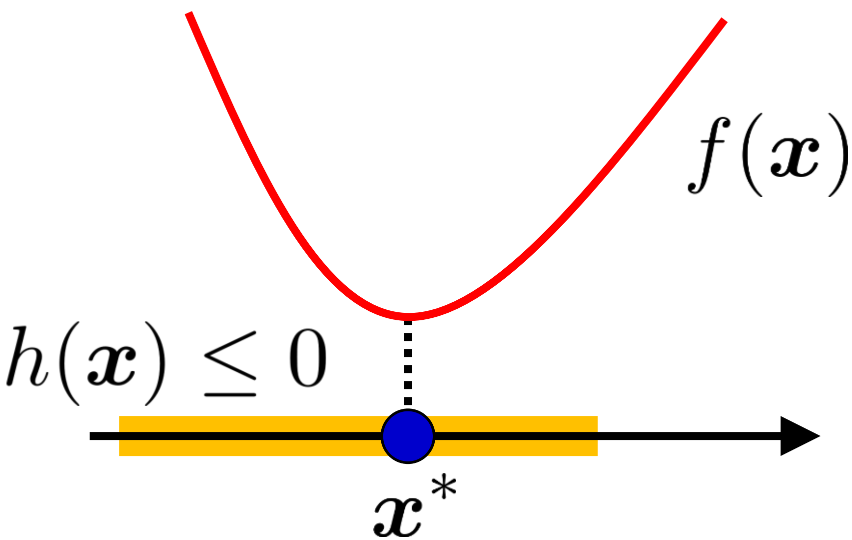
と呼ぶ

最適性条件

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } h(\boldsymbol{x}) \leq 0$$

- 最適解 x^* において制約が有効でないとき，最適解は制約なし最適化問題 $\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x})$ を解くことによって得られる

$$\Rightarrow \nabla f(\boldsymbol{x}^*) = \mathbf{0}$$

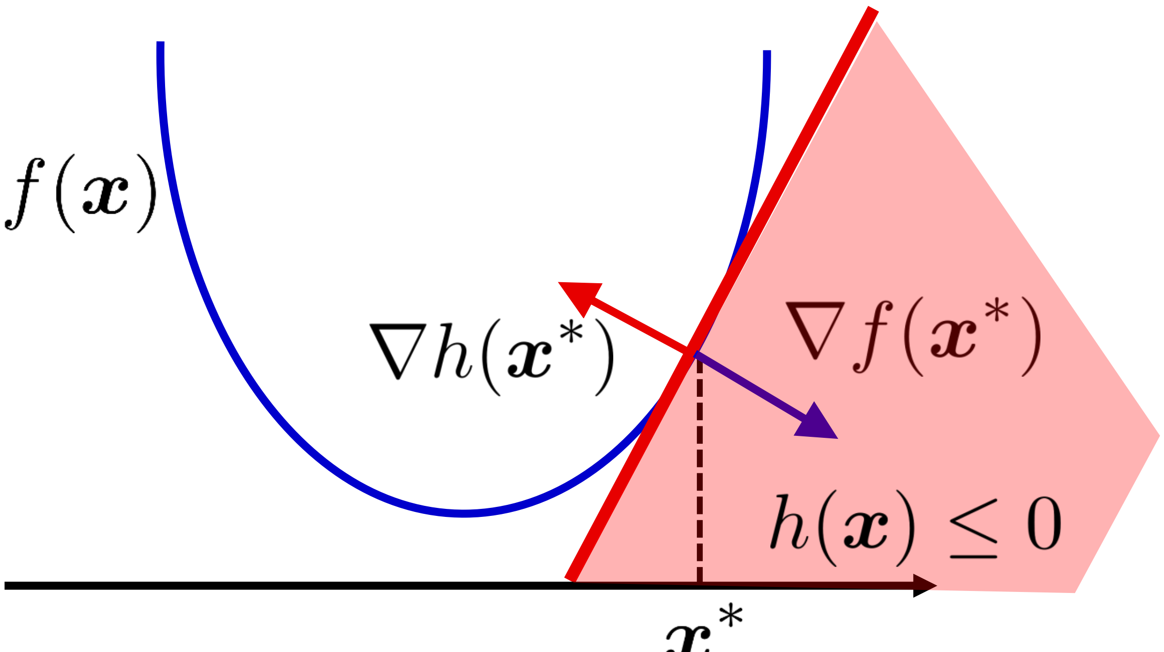


最適性条件

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } h(\boldsymbol{x}) \leq 0$$

- 最適解 x^* において制約が有効なとき，
ラグランジュの未定乗数法より，次式を得る：

$$\nabla f(\boldsymbol{x}^*) + \mu^* \nabla h(\boldsymbol{x}^*) = \mathbf{0} \qquad h(\boldsymbol{x}^*) = 0$$



- また，$\nabla h(\boldsymbol{x}^*)$ の向きより $\mu^* \geq 0$

最適性条件

- まとめると,

$$\nabla f(\boldsymbol{x}^*) + \mu^* \nabla h(\boldsymbol{x}^*) = \mathbf{0}$$

  - $h(\boldsymbol{x}^*) < 0$ ならば $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$ より $\mu^* = 0$
  - $h(\boldsymbol{x}^*) = 0$ ならば $\mu^* \geq 0$

- 更にまとめると

$$\nabla f(\boldsymbol{x}^*) + \mu^* \nabla h(\boldsymbol{x}^*) = \mathbf{0}$$

$$h(\boldsymbol{x}^*) \leq 0$$

$$\mu^* \geq 0$$

$$\mu^* h(\boldsymbol{x}^*) = 0$$

（相補性条件とよぶ.
どちらかは必ずゼロ）

KKT条件

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \mathbf{0},\ \ \boldsymbol{h}(\boldsymbol{x}) \leq \mathbf{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))^\top$$

$$\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), \ldots, h_n(\boldsymbol{x}))^\top$$

- 最適解の必要条件(凸の場合は必要十分条件):
  - $\nabla f(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\top} \nabla \boldsymbol{g}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*\top} \nabla \boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}$
  - $\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}$
  - $\boldsymbol{h}(\boldsymbol{x}^*) \leq \mathbf{0}$
  - $\boldsymbol{\mu}^* \geq \mathbf{0}$
  - $\mu_i^* h_i(\boldsymbol{x}^*) = 0, \quad i = 1, \ldots, n$

カルーシュ・キューン・タッカー
(KKT)条件
(Karush-Kuhn-Tucker conditions)

演習

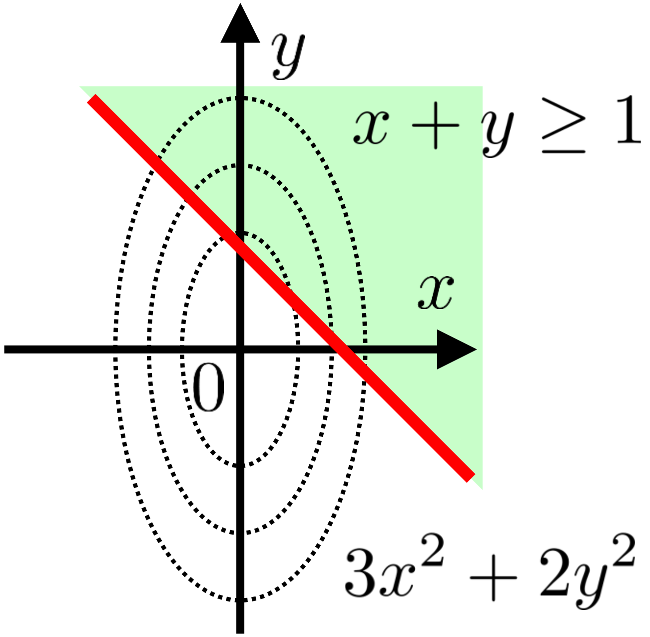
- 以下の最適化問題の解を，KKT条件を用いて求めよ

$$\min_{x,y\in\mathbb{R}} \ 3x^2 + 2y^2$$

$$\text{subject to } x + y \geq 1$$

解答例

$$\min_{x,y\in\mathbb{R}} 3x^2 + 2y^2 \quad \text{subject to } x + y \geq 1$$



- KKT条件は

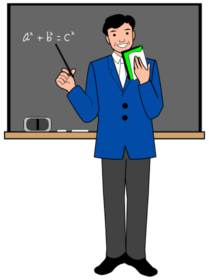
$$6x - \mu = 0$$

$$4y - \mu = 0$$

$$x + y \geq 1$$

$$\mu(1 - x - y) = 0$$

- $\mu = 0$ のとき $x = y = 0$ となり，$x + y \geq 1$ を満たさないので実行可能解でない
- $1 - x - y = 0$ のとき $(x, y) = (0.4, 0.6)$ となり，$x + y \geq 1$を満たすので最適解である



講義の流れ

1. 等式制約付き最適化問題
2. 不等式制約付き最適化問題
   - A) 最適性条件
   - B) 双対問題

双対問題

- 主問題(primal problem):

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0},\ \boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0}$$

$$\boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x}))^\top$$

$$\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), \ldots, h_n(\boldsymbol{x}))^\top$$

- 双対問題(dual problem):
  - 目的関数と制約条件を入れ換えたような形式の問題を考える
  - 制約条件から目的関数を作る
  - 目的関数から制約条件を作る

ラグランジュ双対

$$f^* = \min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0},\ \ \boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0}$$

■ラグランジュ双対問題(Lagrange dual problem):

$$f^* = \max_{\boldsymbol{\lambda},\boldsymbol{\mu}} \omega(\boldsymbol{\lambda},\boldsymbol{\mu})$$

$$\text{subject to } \boldsymbol{\mu} \geq \boldsymbol{0}$$

$$\omega(\boldsymbol{\lambda},\boldsymbol{\mu}) = \inf_{\boldsymbol{x}\in\mathcal{X}} L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})$$

$$L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{\mu}^\top \boldsymbol{h}(\boldsymbol{x})$$

- 凸最適化問題に対しては最適解が一致
- 双対問題の方が制約が単純

演習

- 以下の最適化問題のラグランジュ双対問題を求め，最適解を求めよ

$$\min_{x,y\in\mathbb{R}} \ 3x^2 + 2y^2$$

$$\text{subject to } x + y \geq 1$$

解答例

$$\min_{x,y\in\mathbb{R}} 3x^2 + 2y^2 \quad \text{subject to } x + y \geq 1$$

- $L(x, y, \mu) = 3x^2 + 2y^2 + \mu(1 - x - y)$

$$\frac{\partial L(x,y,\mu)}{\partial x} = 6x - \mu = 0 \quad \Rightarrow \quad x = \frac{\mu}{6}$$

$$\frac{\partial L(x,y,\mu)}{\partial y} = 4y - \mu = 0 \quad \Rightarrow \quad y = \frac{\mu}{4}$$

- $\omega(\mu) = \inf_{x,y} L(x, y, \mu) = -\frac{5}{24}\mu^2 + \mu$

- $\operatorname*{argmax}_{\mu\geq 0} \omega(\mu) = \frac{12}{5} \quad \Rightarrow \quad (x, y) = (0.4, 0.6)$

射影勾配法

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0},\ \boldsymbol{h}(\boldsymbol{x}) \le \boldsymbol{0}$$

双対問題:

$$\max_{\boldsymbol{\lambda},\boldsymbol{\mu}} \omega(\boldsymbol{\lambda},\boldsymbol{\mu}) \quad \text{subject to } \boldsymbol{\mu} \ge \boldsymbol{0}$$

$$\omega(\boldsymbol{\lambda},\boldsymbol{\mu}) = \inf_{\boldsymbol{x}\in\mathcal{X}} L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})$$

$$L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{\mu}^\top \boldsymbol{h}(\boldsymbol{x})$$

射影勾配法:

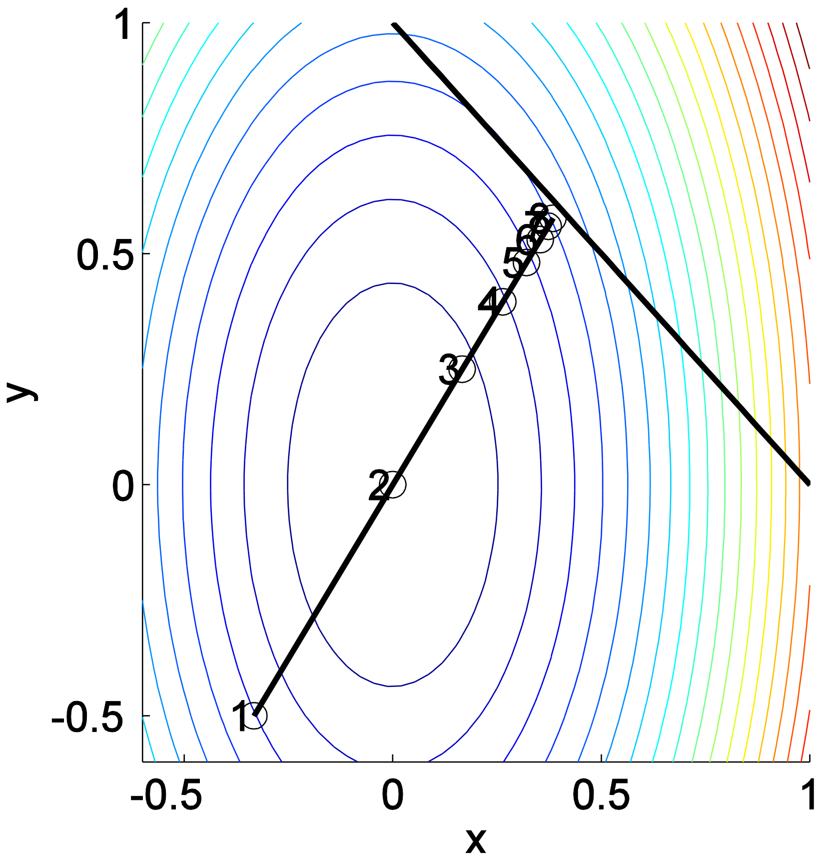
- $\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \varepsilon_k \nabla_{\boldsymbol{\lambda}} \omega(\boldsymbol{\lambda}_k)$
- $\boldsymbol{\mu}_{k+1} = \max(\boldsymbol{0}, \boldsymbol{\mu}_k + \varepsilon_k \nabla_{\boldsymbol{\mu}} \omega(\boldsymbol{\mu}_k))$

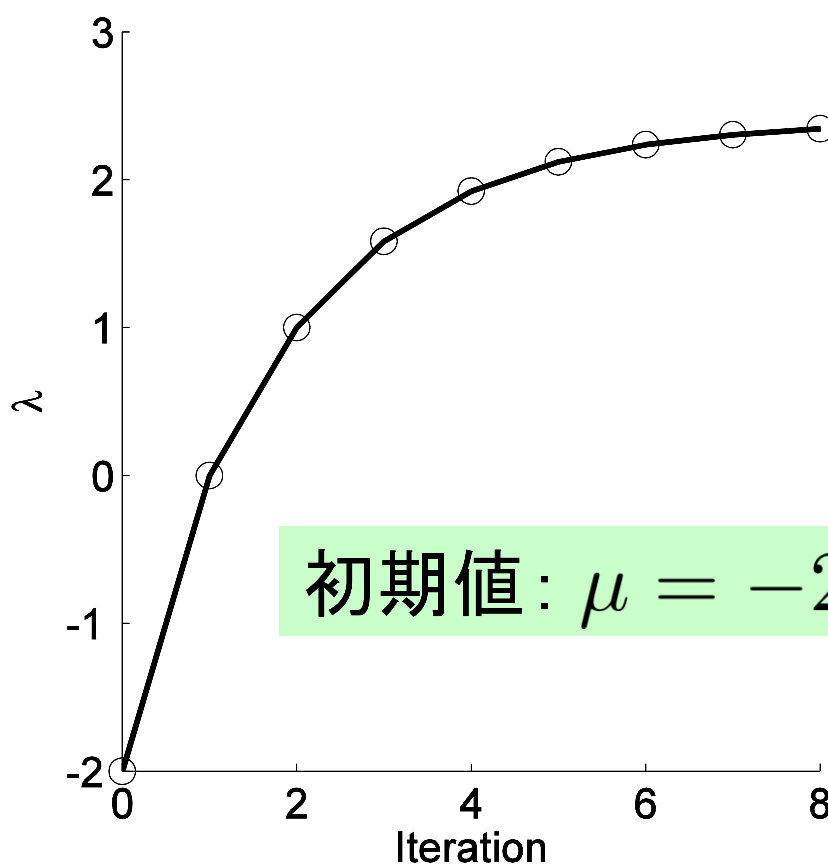
$\varepsilon_k > 0$ ベクトルに対するmaxは要素毎の最大

実行例

$$\min_{x,y\in\mathbb{R}} \ 3x^2 + 2y^2 \quad \text{subject to } x + y \geq 1$$

$$\omega(\mu) = -\frac{5}{24}\mu^2 + \mu$$





初期値: $\mu = -2$

まとめ

■ 制約付き最適化問題:

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0},\ \ \boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0}$$

- ラグランジュ関数:

$$L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^\top \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{\mu}^\top \boldsymbol{h}(\boldsymbol{x})$$

- 双対問題:

$$\max_{\boldsymbol{\lambda},\boldsymbol{\mu}} \inf_{\boldsymbol{x}\in\mathcal{X}} L(\boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \quad \text{subject to } \boldsymbol{\mu} \geq \boldsymbol{0}$$

- KKT条件:

$$\nabla_{\boldsymbol{x}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \boldsymbol{0}$$

$$\nabla_{\boldsymbol{\lambda}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \boldsymbol{0} \qquad \boldsymbol{\mu}^* \geq \boldsymbol{0}$$

$$\nabla_{\boldsymbol{\mu}} L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \leq \boldsymbol{0} \qquad \mu_i^* h_i(\boldsymbol{x}^*) = 0, \quad \forall i$$

宿題1

- 二次計画(quadratic programming)問題

$$\min_{\boldsymbol{x}} \frac{1}{2}\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} \quad \text{subject to } \boldsymbol{C}\boldsymbol{x} = \boldsymbol{d}$$
$$\boldsymbol{E}\boldsymbol{x} \leq \boldsymbol{f}$$

  のKKT条件を求めよ
  - ベクトルの不等式は要素毎の不等式

- $\boldsymbol{A} = \boldsymbol{O}$のとき，線形計画(linear programming)問題とよぶ

宿題2

- 障壁法と罰則法を実装し，以下の最適化問題の解を数値的に求めよ:

$$\min_{x,y\in\mathbb{R}} \ 3x^2+2y^2 \quad \text{subject to } x+y\geq 1$$

- **障壁法**:

$$\boldsymbol{x}_{k+1}=\operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x})-c_k\log(-h(\boldsymbol{x}))$$

$$c_1>c_2>\cdots>0$$

- **罰則法**:

$$\boldsymbol{x}_{k+1}=\operatorname*{argmin}_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x})+c_k\max(0,h(\boldsymbol{x}))^2$$

$$0<c_1<c_2<\cdots$$

$$\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \quad \text{subject to } h(\boldsymbol{x})\leq 0$$