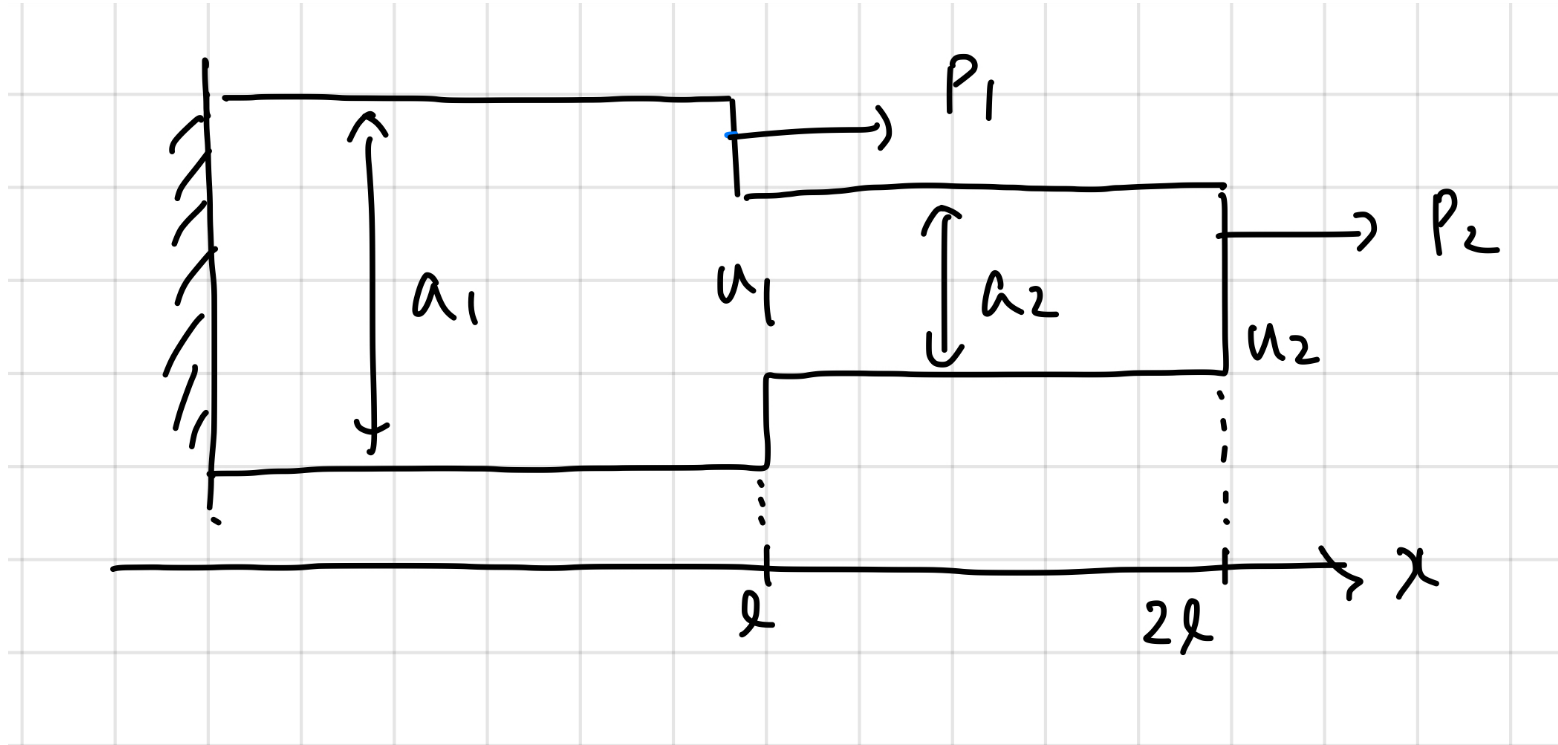
# 力学的シンセシス (4)

システムデザイン工学科
**飯盛 浩司**

# はじめに

**[解きたい問題] 力学的シンセシスで取り扱う問題の例：**

Find $\boldsymbol{a} = (a_1, a_2)^t$
such that $\min\ f(\boldsymbol{a}) = \boldsymbol{u} \cdot \boldsymbol{p}$
subject to $K\boldsymbol{u} = \boldsymbol{p}$
$a_1\ell + a_2\ell \leq c$




前回, 随伴変数法を学び, 等式制約を満たす設計変数に対して $\nabla f$ を計算できるようになった. →不等式制約さえ別で考慮すれば良い. 以降,

Find $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^t$
such that $\min\ f(\boldsymbol{x})$ subject to $g_i(\boldsymbol{x}) \leq 0$ (for $i = 1, \cdots, m$)

の数値計算法を学ぶ. **今日の目標: 制約無し最適化問題の数値解法**を知る.

# 反復法

局所最適解を探すためには $\underline{\nabla f(\boldsymbol{x}) = \boldsymbol{0}}$ を解けば良い.
↑ 一般に非線形方程式.
ほとんどの場合(解析的には)解けない.

## 反復法

0. $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$ を与える. $k = 0$ とする.

1. 終了条件を満たすなら $\boldsymbol{x}^{(k)}$ を最適解として出力.

2. **探索方向** $\boldsymbol{d}^{(k)} \in \mathbb{R}^n$ とステップサイズ $\alpha^{(k)} > 0$ を何らかの方法で計算.

3. $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}$ と設計変数を更新. $k \leftarrow k + 1$ として1.に戻る.

終了条件の例: $\|\nabla f(\boldsymbol{x}^{(k)})\|$ が十分に小さい.
$\alpha^{(k)}$ が十分に小さい.
$k$ が十分に大きい.
$f(\boldsymbol{x}^{(k)})$ が十分に小さい.

# 探索方向の決め方: 降下法

**降下法**: $f(\boldsymbol{x}^{(0)}) > f(\boldsymbol{x}^{(1)}) > f(\boldsymbol{x}^{(2)}) > \cdots$ となるように $\boldsymbol{x}^{(k)}$ を生成

ステップサイズ $\alpha^{(k)} > 0$ が十分に小さければ

$$\begin{aligned} f(\boldsymbol{x}^{(k+1)}) &= f(\boldsymbol{x}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}) \\ &\simeq f(\boldsymbol{x}^{(k)}) + \alpha^{(k)}\nabla f(\boldsymbol{x}^{k}) \cdot \boldsymbol{d}^{(k)} \end{aligned}$$

なので, $\nabla f(\boldsymbol{x}^{(k)}) \cdot \boldsymbol{d}^{(k)} < 0$ を満たすような探索方向を選べばよい.

# 最急降下法

## 最急降下法

$\boldsymbol{d}^{(k)} = -\nabla f(\boldsymbol{x}^{(k)})$ と探索方向を選ぶ.

$\nabla f(\boldsymbol{x}^{(k)}) \cdot \boldsymbol{d}^{(k)} = -\|\nabla f(\boldsymbol{x}^{(k)})\|^2 < 0$ なので, 常に降下方向を生成できる.

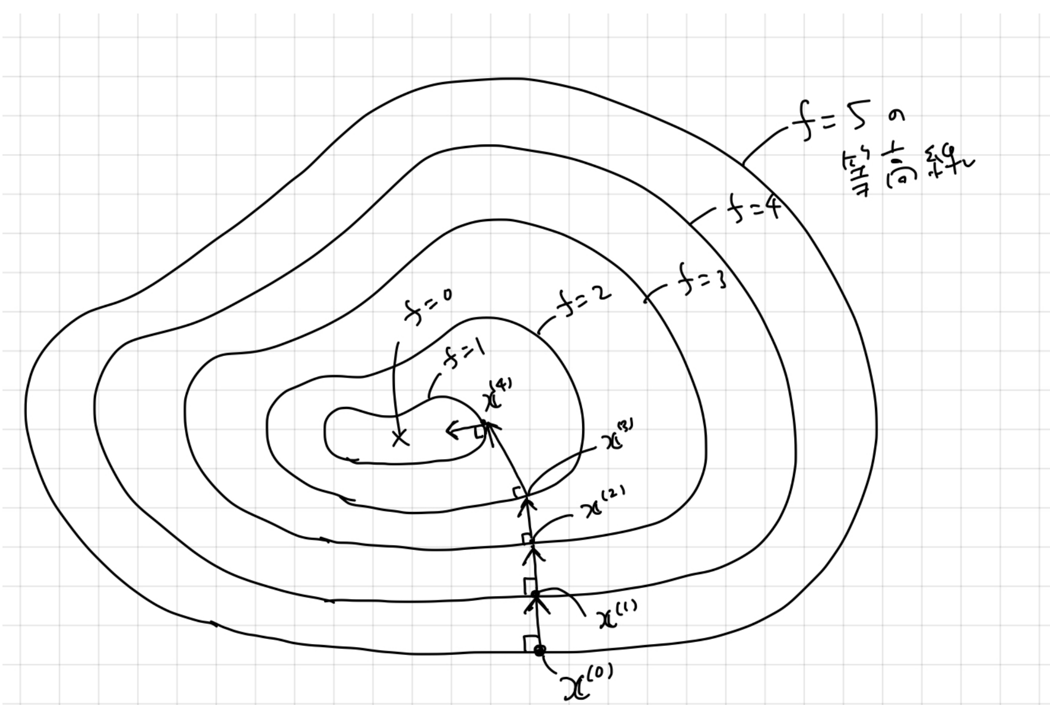



図1. 最急降下法のイメージ

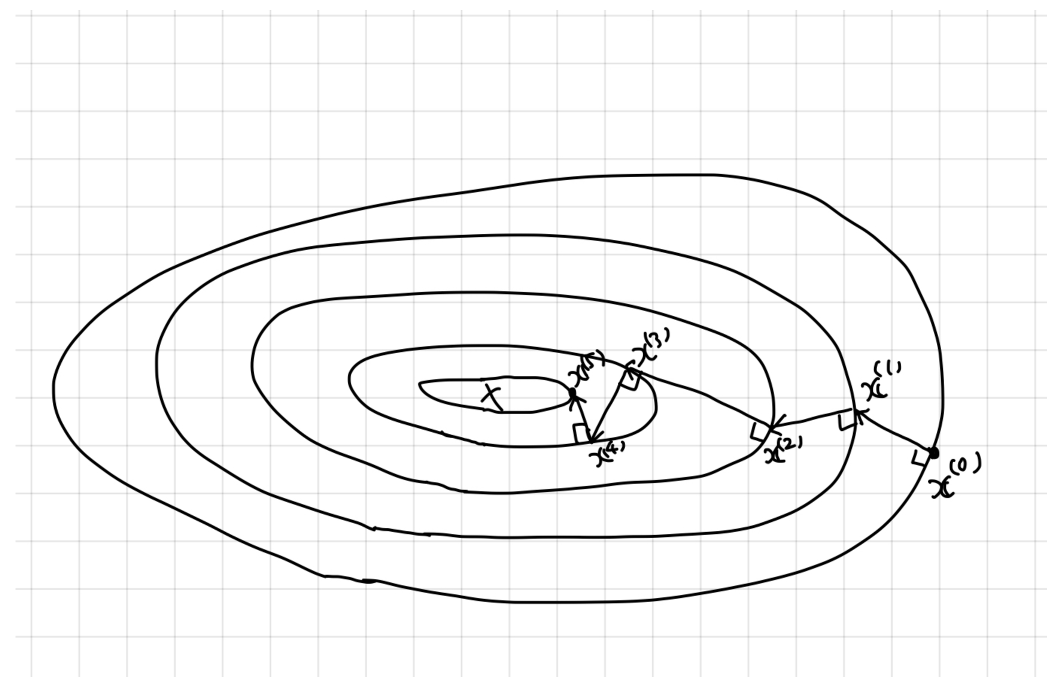


図2. 最急降下法は遅い(ことがある)

**Note:** 収束は一般に遅いが, 局所最適解への収束は保証される.

# Newton法

## Newton法

$\boldsymbol{d}^{(k)} = -H_f^{-1}(\boldsymbol{x}^{(k)})f(\boldsymbol{x}^{(k)})$ と探索方向を選ぶ.

(設計の指針) $\nabla f(\boldsymbol{x}^{(k+1)}) = 0$ を目指す.

ステップサイズ $\alpha^{(k)} > 0$ が十分に小さければ

$$\begin{aligned}\nabla f(\boldsymbol{x}^{(k+1)}) &= \nabla f(\boldsymbol{x}^{(k)} + \alpha^{(k)}\boldsymbol{d}^{(k)}) \\ &\simeq \nabla f(\boldsymbol{x}^{(k)}) + \alpha^{(k)}H_f(\boldsymbol{x}^{k})\boldsymbol{d}^{(k)}\end{aligned}$$

$\Rightarrow$ もし $\nabla f(\boldsymbol{x}^{(k+1)}) = \mathbf{0}$ ならば $\nabla f(\boldsymbol{x}^{(k)}) + \alpha^{(k)}H_f(\boldsymbol{x}^{(k)})\boldsymbol{d}^{(k)} = \mathbf{0}$

$\Rightarrow$ 探索方向として, $\boldsymbol{d}^{(k)} = -H_f^{-1}(\boldsymbol{x}^{(k)})\nabla f(\boldsymbol{x}^{(k)})$ を使ってみよう.

**Note:** $H_f^{-1}(\boldsymbol{x}^{(k)})$ が正定値でなければ**降下法ではない**.

Hesse行列とその逆の計算が大変.

初期値を上手に選べば収束は速いが最適解への収束は保証されない.

# 準Newton法 1/2

## Newton法

$$\boldsymbol{d}^{(k)} = -H_f^{-1}(\boldsymbol{x}^{(k)})f(\boldsymbol{x}^{(k)})$$ と探索方向を選ぶ.

## 最急降下法

$$\boldsymbol{d}^{(k)} = -\nabla f(\boldsymbol{x}^{(k)})$$ と探索方向を選ぶ.

↑比較すると, $I$ (単位行列)を $H_f^{-1}(\boldsymbol{x}^{(k)})$ に置き換えただけに見える.

収束性を保証しつつそこそこ速い行列の選び方はない？

## 準Newton法

$$\boldsymbol{d}^{(k)} = -\mathrm{B}_{(k)}^{-1}\nabla f(\boldsymbol{x}^{(k)})$$ と探索方向を選ぶ.

# 準Newton法 2/2

$\mathrm{B}_{(k)}$ の設計指針:

- $\mathrm{B}_{(k)}$ は対称行列が良い.
- $\mathrm{B}_{(k)}$ は正定値行列が良い.
- $\mathrm{B}_{(k)}$ とその逆を簡単に(=小さな計算負荷で)計算したい.
- $\mathrm{B}_{(k)}$ は, Hesse行列を(何らかの意味で)近似して欲しい.
- $\mathrm{B}_{(k)}$ は漸化的に決めれば良い(初期値は $I$ とか).

## BFGS公式

$$\mathrm{B}_{(k+1)} = \mathrm{B}_{(k)} - \frac{\mathrm{B}_{(\mathrm{k})}\boldsymbol{s}_{(k)}(\mathrm{B}_{(\mathrm{k})}\boldsymbol{s}_{(k)})^t}{\boldsymbol{s}^t_{(k)}\mathrm{B}_{(k)}\boldsymbol{s}_{(k)}} + \frac{\boldsymbol{y}_{(k)}\boldsymbol{y}^t_{(k)}}{\boldsymbol{s}^t_{(k)}\boldsymbol{y}_{(k)}}$$

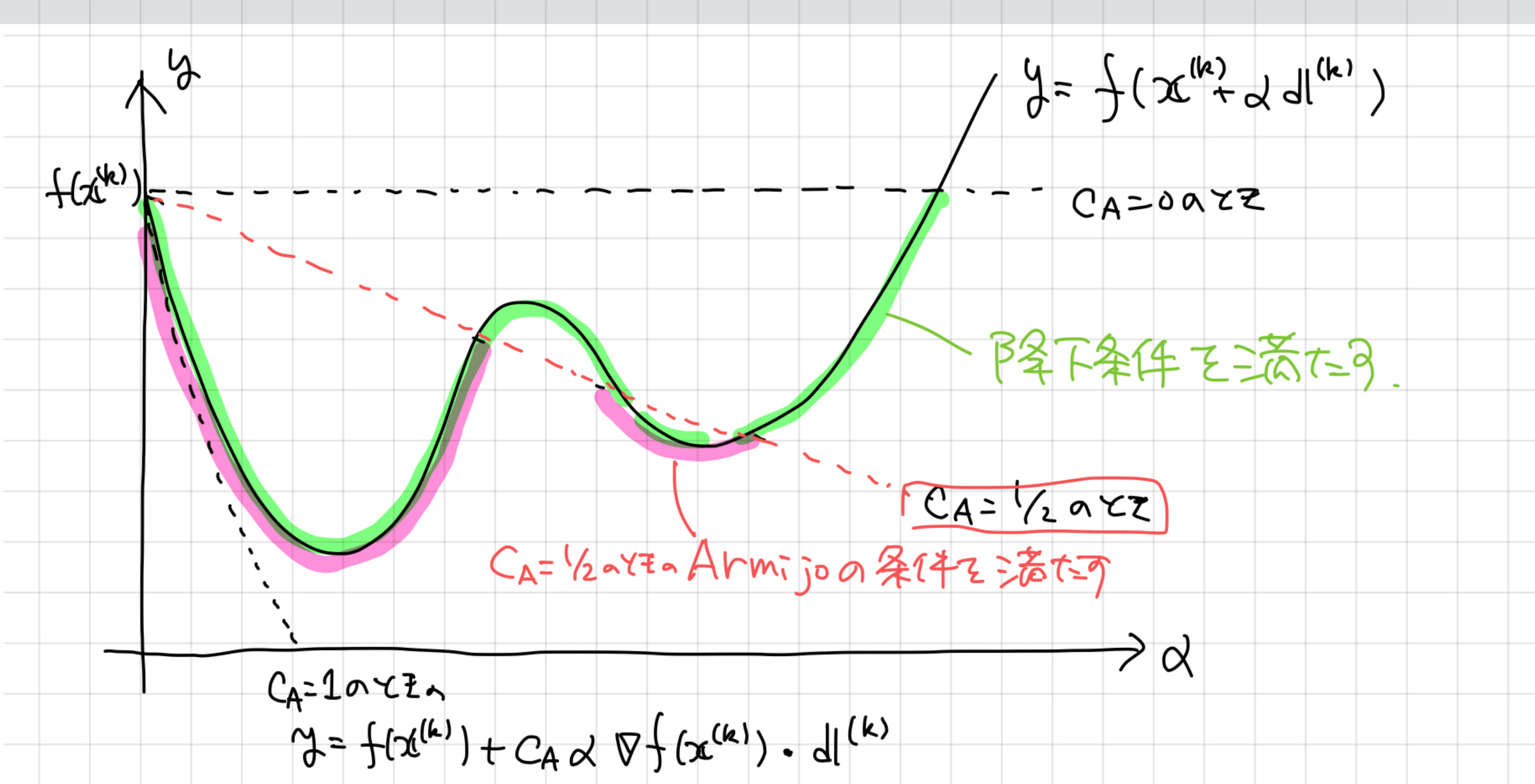
ただし, $\boldsymbol{s}_{(k)} = \boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)},\ \ \boldsymbol{s}_{(k)} = \nabla f(\boldsymbol{x}^{(k+1)}) - \nabla f(\boldsymbol{x}^{(k)})$.

# ステップサイズの決め方 1/2

## Armijoのルール … 降下条件を十分に満たして欲しい

$$f(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}) \leq f(\boldsymbol{x}^{(k)}) + C_A \alpha \nabla f(\boldsymbol{x}^{(k)}) \cdot \boldsymbol{d}^{(k)} \ ( < f(\boldsymbol{x}^{(k)}))$$

を満たすように $\alpha^{(k)}$ を選択する. ここに, $0 < C_A < 1$ はパラメータ.



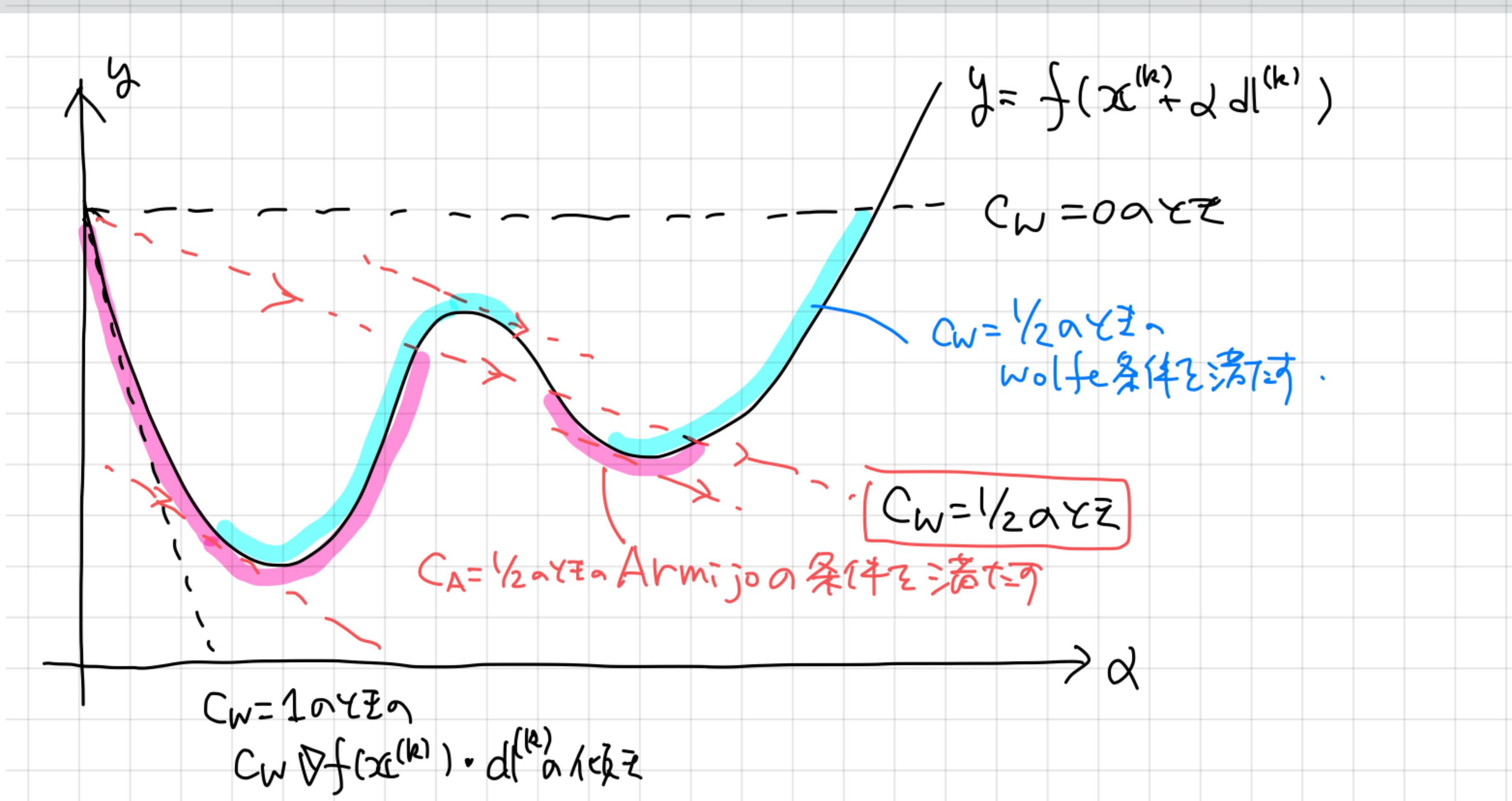

**Note:** 通常, $C_A$ は小さな値 (0.001とか) を用いる.
Armijoのルールだけだと, $\alpha$ はどんどん小さくなる傾向がある.

# ステップサイズの決め方 2/2

## Wolfeのルール … ステップサイズがあまり小さくならないで欲しい

$$\nabla f(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}) \cdot \boldsymbol{d}^{(k)} \leq C_W \nabla f(\boldsymbol{x}^{(k)}) \cdot \boldsymbol{d}^{(k)}$$

を満たすように $\alpha^{(k)}$ を選択する. ここに, $C_A < C_W < 1$ はパラメータ.




**Note:** ArmijoのルールとWolfeのルールの両方を満たすようにステップサイズを決めると良さそう.

# Quiz: Why $c_A < c_W$ ?

# まとめ

## 反復法

0. $\boldsymbol{x}^{(0)} \in \mathbb{R}^n,\ 0 < C_A < C_W < 1,\ \varepsilon > 0,\ \alpha^{(0)} > 0$ を与える. $k = 0$ とする.

1. 終了条件を満たすなら $\boldsymbol{x}^{(k)}$ を最適解として出力.

2. **探索方向** $\boldsymbol{d}^{(k)} \in \mathbb{R}^n$ を最急降下法や(準)Newton法で計算.

3. $\alpha^{(k)}$ が**Armijoのルール** $f(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}) \le f(\boldsymbol{x}^{(k)}) + C_A \alpha \nabla f(\boldsymbol{x}^{(k)}) \cdot \boldsymbol{d}^{(k)}\ (< f(\boldsymbol{x}^{(k)}))$ を満たさないなら, 満たすまで $\alpha^{(k)} \leftarrow \alpha^{(k)}/(1+\varepsilon)$ とする (減速).

4. $\alpha^{(k)}$ が**Wolfeのルール** $\nabla f(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}) \cdot \boldsymbol{d}^{(k)} \le C_W \nabla f(\boldsymbol{x}^{(k)}) \cdot \boldsymbol{d}^{(k)}$を満たさないなら, $\alpha^{(k)} \leftarrow \alpha^{(k)}/(1-\varepsilon)$ として, 2に戻る (加速).

5. $\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha^{(k)} \boldsymbol{d}^{(k)}$ と設計変数を更新. $k \leftarrow k+1$ として1.に戻る.