

数値解析特論 課題 2

提出日：2023 年 8 月 1 日

M213325

斉藤正裕

1.入力画像を次の相似変換によって変換した出力画像から回転角度 θ とスケールパラメータ s をガウス・ニュートン法によって推定しなさい。

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = s \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

ガウス・ニュートン法は以下のように最小化を行う。

① θ と s に任意の初期値を与える。入力画像 $I(x,y)$ と出力画像 $I'(x',y')$ としたとき、次式が最小となるパラメータを推定する。

$$J = \frac{1}{2} \sum_{(x,y) \in I} (I'(x',y') - I(x,y))^2$$

②出力画像 I' に対して平滑微分画像 I'_x 、 I'_y を作成する。平滑微分画像はガウシアンフィルタで画像のノイズを低減し、その画像に対してx軸方向、y軸方向でそれぞれ微分することで得ることができる。

③ $J = \frac{1}{2} \sum_{(x,y) \in I} (I'(x',y') - I(x,y))^2$ の θ に対する1階微分 J_θ と2階微分 $J_{\theta\theta}$ を計算する。

$$J_\theta = \sum_{(x,y) \in I} (I'(x',y') - I(x,y)) \left(I'_{x'} \frac{dx'}{d\theta} + I'_{y'} \frac{dy'}{d\theta} \right)$$

$$J_{\theta\theta} = \sum_{(x,y) \in I} \left(I'_{x'} \frac{dx'}{d\theta} + I'_{y'} \frac{dy'}{d\theta} \right)^2$$

このとき、

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = s \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

と置くと $\frac{dx'}{d\theta}$ 、 $\frac{dy'}{d\theta}$ は次の式で求められる。

$$\frac{dx'}{d\theta} = \sum_{i,j=1,2} \frac{\partial x'}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial \theta}, \quad \frac{dy'}{d\theta} = \sum_{i,j=1,2} \frac{\partial y'}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial \theta}$$

④ J の s に対する1階微分 J_s と2階微分 J_{ss} を計算する。

$$J_s = \sum_{(x,y) \in I} (I'(x', y') - I(x, y)) \left(I'_{x'} \frac{dx'}{ds} + I'_{y'} \frac{dy'}{ds} \right)$$
$$J_{ss} = \sum_{(x,y) \in I} \left(I'_{x'} \frac{dx'}{ds} + I'_{y'} \frac{dy'}{ds} \right)^2$$

また、 $\frac{dx'}{ds}$ 、 $\frac{dy'}{ds}$ は次の式で求められる。

$$\frac{dx'}{ds} = \sum_{i,j=1,2} \frac{\partial x'}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial s}, \quad \frac{dy'}{ds} = \sum_{i,j=1,2} \frac{\partial y'}{\partial a_{ij}} \frac{\partial a_{ij}}{\partial s}$$

⑤ J の θ と s での微分 $J_{\theta s}$ を計算する。

$$J_{\theta s} = \sum_{(x,y) \in I} \left(I'_{x'} \frac{dx'}{d\theta} + I'_{y'} \frac{dy'}{d\theta} \right) \left(I'_{x'} \frac{dx'}{ds} + I'_{y'} \frac{dy'}{ds} \right)$$

⑥ $(\Delta\theta, \Delta s)$ を次のように計算する。

$$\begin{pmatrix} \Delta\theta \\ \Delta s \end{pmatrix} = - \begin{pmatrix} J_{\theta\theta} & J_{\theta s} \\ J_{\theta s} & J_{ss} \end{pmatrix}^{-1} \begin{pmatrix} J_{\theta} \\ J_s \end{pmatrix}$$

⑦ $\theta \leftarrow \theta + \Delta\theta$ 、 $s \leftarrow s + \Delta s$ として収束するまで繰り返す。

実験内容

今回の実験では以下の3つの観点で実験を行った。

1. 初期値を変えたときの収束の仕方の違い
2. 入力画像を変えたときの推定精度の違い
3. ガウシアンフィルタのカーネルサイズによる精度と収束の違い

はじめに入力画像を図1(a)、($\theta=45, s=0.5$)で相似変換した出力画像を図1(b)とする。どちらも画素数は 270×270 である。



(a)入力画像

(b)出力画像

図1 実験に使用した画像

これらの画像を使い、 θ と s の推定を行った。本実験で使用した値を以下に示す。

- ・ θ の初期値：0
- ・ s の初期値：1
- ・ ループ回数：100
- ・ ガウシアンフィルタのカーネルサイズ： 7×7

ループごとの θ の推定値をグラフ化したものを図 2、 s の推定値をグラフ化したものを図 3 に示す。

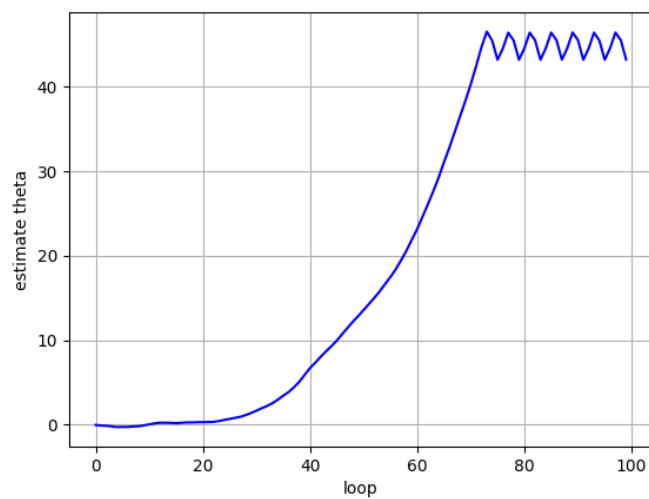


図 2 θ の推定値の収束グラフ

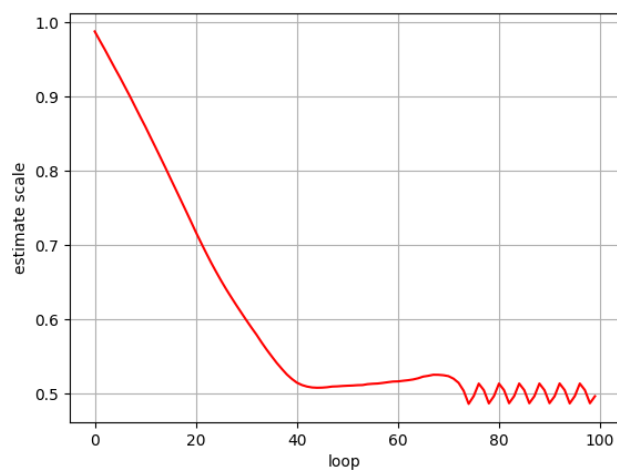


図 3 s の推定値の収束グラフ

最終的に $\theta = 43.2458$ 、 $s = 0.4963$ となり、真値に非常に近くなっている。しかし、グラフを見ると、どちらも 70 回以降で収束せず振動していることが分かる。

1. 初期値を変えたときの収束の仕方の違い

次に θ の初期値を 20、 s の初期値を 0.8 としたときの収束の仕方を調べる。ループごとの θ の推定値をグラフ化したものを図 4、 s の推定値をグラフ化したものを図 5 に示す。

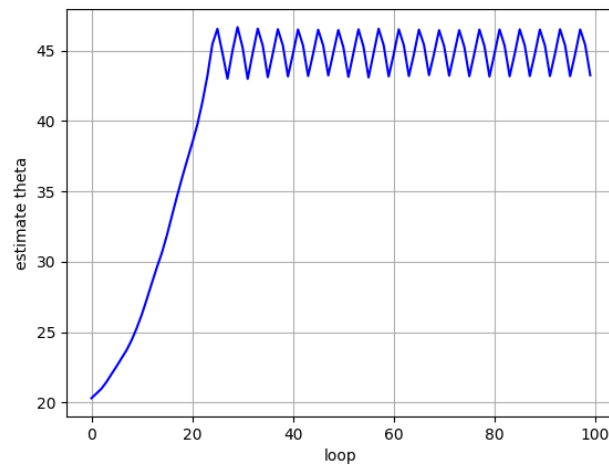


図 4 初期値を変更時の θ の推定値の収束グラフ

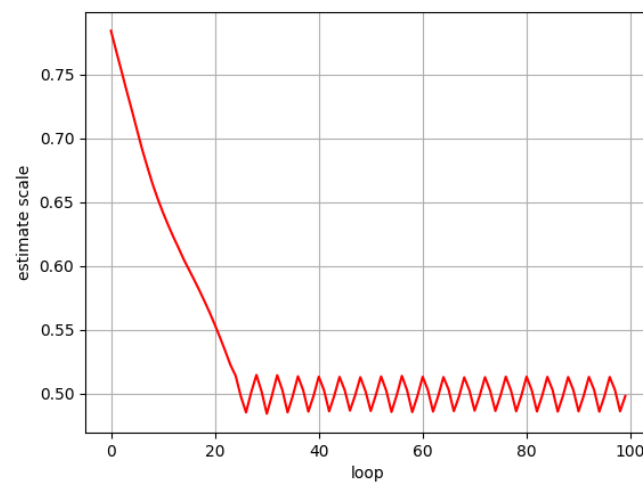
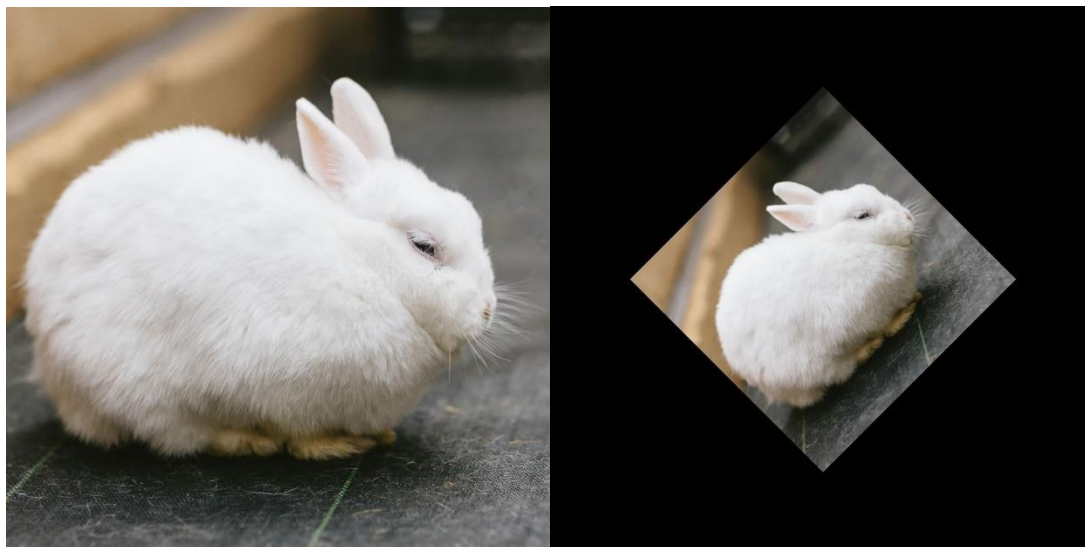


図 5 初期値を変更時の s の推定値の収束グラフ

推定結果は $\theta = 43.2503$ 、 $s = 0.4982$ となった。どちらのグラフでも初期値を真値に近づけたことで、少ないループ回数で真値付近まで近づいたが、その後収束することなく振動した。従って、初期値は真値に近づく回数にのみ影響していると分かる。

2. 入力画像を変えたときの推定精度の違い

次に入力画像とし出力画像を画素数の高さとの向き取りやすさに着目し、画像に変えて実験した。使用した画像を図6に示す。 $\theta=45$ 、 $s=0.5$ で相似変換した。画素数はどちらの画像も 953×953 である。



(a)入力画像

(b)出力画像

図6 実験に使用した画素数の高い画像

ループごとの θ の推定値をグラフ化したものを図7、 s の推定値をグラフ化したものを図8に示す。

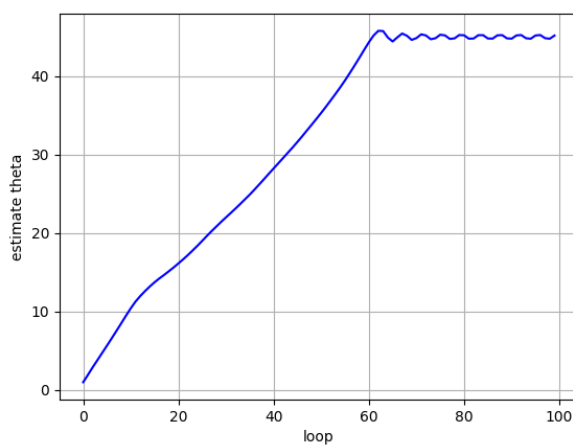


図7 画像変更時の θ の推定値の収束グラフ

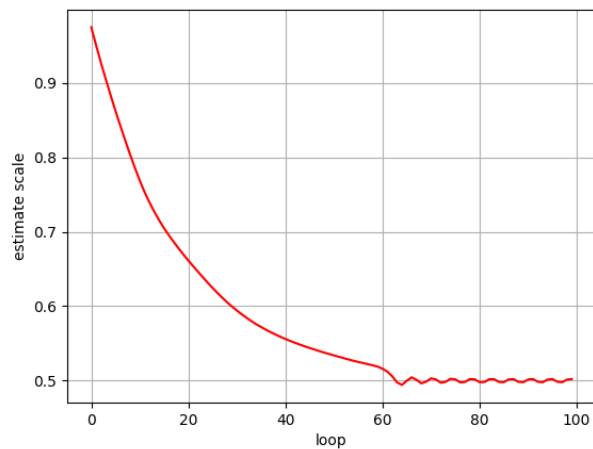


図 8 画像変更時の s の推定値の収束グラフ

最終的な推定結果は $\theta = 45.1672$ 、 $s = 0.5022$ となった。図 7、図 8 両方とも画像を変えたことで真値まで近づく回数が少なくなり、振動の振幅も小さくなっている。この影響により、最終的な推定結果も真値に近づいていることが分かる。従って、画素数を上げることで得られる情報が多くなり、振動を低減することができる。

3. ガウシアンフィルタのカーネルサイズによる精度と収束の違い

最後にガウシアンフィルタのカーネルサイズを 5×5 に変更したときのループごとの、 θ の推定値をグラフ化したものを図 9、 s の推定値をグラフ化したものを図 10 に示す。

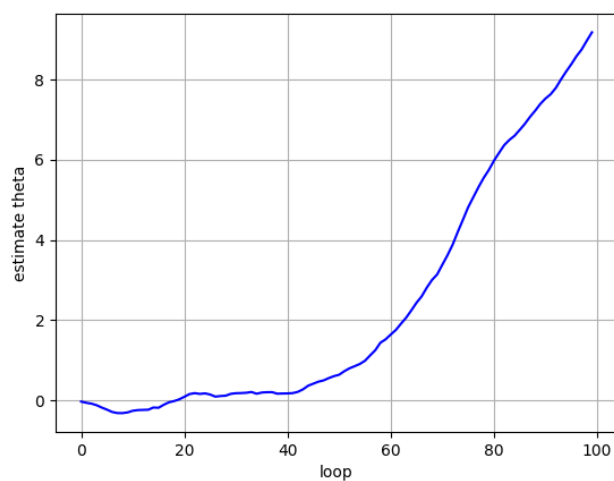


図 9 カーネルサイズ変更時の θ の推定値の収束グラフ

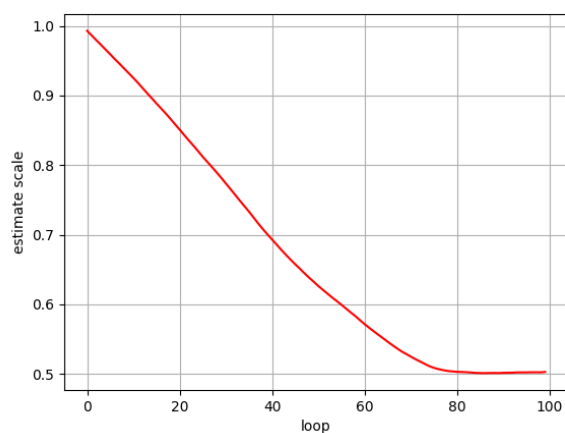


図 10 カーネルサイズ変更時の s の推定値の収束グラフ

推定結果は $\theta = 9.1752$ 、 $s = 0.5028$ となった。グラフを見ると振動せず収束しており、 s の最終結果は真値に非常に近い値になった。しかし、 θ の最終結果は真値から大きくずれている。画素数に合わせて適切なバッチサイズに合わせることで推定結果を収束させることができるが、平滑微分画像にノイズが多く含まれてしまい、 θ に大きな誤差が生じてしまうと分かった。

以上の結果より、画素数な大きな画像に対し、適切なカーネルサイズで平滑化することで、振動を軽減してパラメータの推定が行えるといえる。また、初期値を真値に近づけることで真値まで近づく推定回数を大幅に少なくすることができる。