

Evaluation the stability of Bézier fitting based algorithm for multi-objective optimization problem

Representation Learning Seminar

Yasunari HIKIMA

September 8, 2021

Topic

- Proposed algorithm
- Stability of algorithm
- Numerical experiments

A multi-objective optimization problem

We consider the following multi-objective optimization problem:

$$\begin{aligned} \min_x \quad & f(\mathbf{x}) := (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_M(\mathbf{x})) , \\ \text{s.t.} \quad & \mathbf{x} \in X \subseteq \mathbb{R}^n . \end{aligned}$$

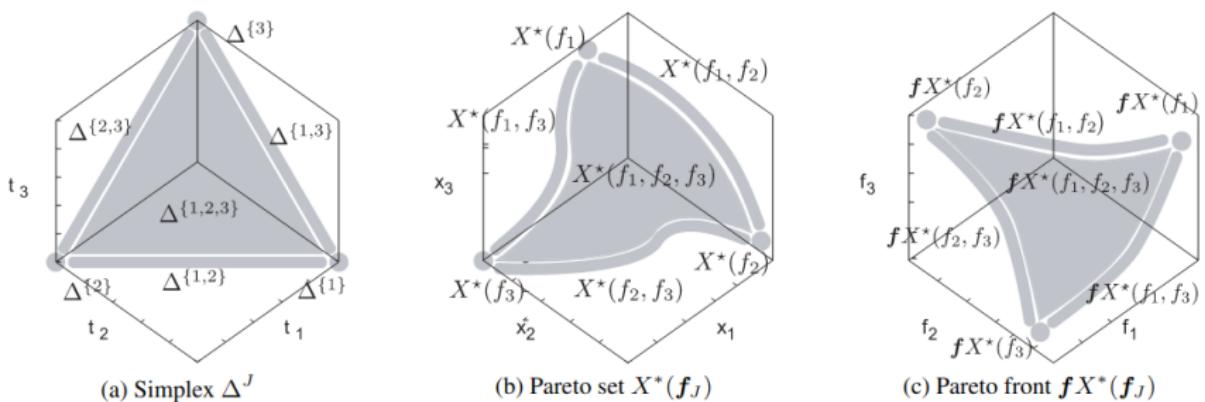


Figure 1: Skeletons for non-empty $J \subseteq \{ 1, 2, 3 \}$ and simplicial $\mathbf{f} = (f_1, f_2, f_3)$.

Cite: [Kobayashi+ 2019]

Algorithm

The proposed algorithm is described as follows:

- ① 初期の制御点 $\mathbf{p}^{(1)} \in \mathbb{R}^D$ を設定する. $k \leftarrow 1$ とする.
- ② 標準単体 Δ^{M-1} 上から点集合 $\{\mathbf{t}_i\}_{i=1}^N \subseteq \mathbb{R}^M$ を生成する. 制御点 $\mathbf{p}^{(k)}$ で表現されるベジエ単体の式に代入して, 点集合 $\{\mathbf{x}_i^{(k)}\}_{i=1}^N \subseteq \mathbb{R}^n$ を得る.
- ③ 点集合 $\{\mathbf{x}_i^{(k)}\}_{i=1}^N$ を以下の規則によって更新する:
$$\mathbf{x}_i^{(k+1)}(\mathbf{t}) = \mathbf{b}^{(k)}(\mathbf{t}_i | \mathbf{p}^{(k)}) + \alpha d_x \tilde{f}(\mathbf{b}^{(k)}(\mathbf{t}_i | \mathbf{p}^{(k)})) \quad (i = 1, \dots, N).$$
- ④ 制御点を以下の規則によって更新する:

$$\mathbf{p}^{(k+1)} = \arg \min_{\mathbf{p} \in \mathbb{R}^D} \int_{\Delta^{M-1}} \|\mathbf{x}^{(k+1)}(\mathbf{t}) - \mathbf{b}(\mathbf{t} | \mathbf{p})\|^2 d\mathbf{t}.$$

- ⑤ 終了条件を満たしているか確認し, 満たしていれば $\mathbf{p}^{(k+1)}$ を出力して反復を終了する. さもなければ, $k \leftarrow k + 1$ として step2 へ戻る.

Some remarks on the algorithm I

- 各種パラメータは以下の通り:
 - M : 解くべき多目的最適化問題の目的関数の個数
 - n : 解くべき多目的最適化問題の決定変数の次元
 - N : アルゴリズムで用いるサンプル点の個数
 - D : (求める) ベジエ単体の次数
- 写像 $b: \Delta^{M-1} \ni t \mapsto b(t) \in \mathbb{R}^M$ は次式で定義される:

$$b(t) := \sum_{d \in \mathbb{N}_D^M} \binom{D}{d} t^d p_d,$$

where

$$\binom{D}{d} := \frac{D!}{d_1! d_2! \dots d_M!}, \quad t^d := t_1^{d_1} t_2^{d_2} \dots t_M^{d_M},$$

$$\mathbb{N}_D^M := \left\{ \mathbf{d} = (d_1, d_2, \dots, d_M) \in \mathbb{N}^M \mid \sum_{m=1}^M d_m = D \right\}$$

Some remarks on the algorithm II

- 各 i に対してベクトル z_i を

$$z_i := \left[\binom{D}{d_1} t_{i,1}^{\mathbf{d}_1}, \binom{D}{d_2} t_{i,2}^{\mathbf{d}_2}, \dots, \binom{D}{d_{|\mathbb{N}_D^M|}} t_{i,|\mathbb{N}_D^M|}^{\mathbf{d}_{|\mathbb{N}_D^M|}} \right]^\top \in \mathbb{R}^{|\mathbb{N}_D^M|}$$

で定義する。また、行列 $P \in \mathbb{R}^{|\mathbb{N}_D^M| \times M}$ と行列 $Z \in N \times |\mathbb{N}_D^M|$ を

$$P := [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{|\mathbb{N}_D^M|}], \quad Z = [z_1, z_2, \dots, z_N]$$

で定義すると写像 b は $b(t) = ZP$ で表される。

- 点 x の更新式に現れる関数 \tilde{f} は M 個の目的関数を t で重み付けをした関数を表す, i.e.,

$$\tilde{f}(\mathbf{x}) := \sum_{i=1}^N t_i f_i(\mathbf{x}).$$

A stability of an algorithm

Def. Uniform stability

乱択アルゴリズム A が ϵ -uniformly stable であるとは，高々一つのデータしか異なるような任意のサンプル集合 $S, S' \in Z^n$ に対して，以下の不等式が成り立つことをいう：

$$\sup_z \mathbb{E}_A [|f(A(S); z) - f(A(S'); z)|] \leq \epsilon.$$

ここで， f は誤差関数を表す．

提案アルゴリズムに対する安定性

提案アルゴリズム A が ϵ -uniformly stable であるとは，高々一つのデータしか異なるような任意のサンプル集合 $S, S' \in Z^n$ に対して，以下の不等式が成り立つことをいう：

$$|\text{loss}(A(S); z) - \text{loss}(A(S'); z)| \leq \epsilon.$$

An upper bound of loss function

誤差関数

提案手法における誤差関数は次式で定義される:

$$\text{loss}(P; t) := \|b(t|P) - b^*(t)\|.$$

ここで, t は標準単体上の点, P は制御点を並べた行列, $b(\cdot|P)$ は制御点行列 P のもとでのベジエ単体, b^* は真のベジエ単体をそれぞれ表す.

ベジエ単体の表現

ベジエ写像 $b: \Delta^{M-1} \rightarrow \mathbb{R}^M$ は次式で表される:

$$b(t) := \sum_{d \in \mathbb{N}_D^M} \binom{D}{d} t^d p_d = P^\top z,$$

ここで, ベクトル z は係数を並べた行列である.

Derivation of the upper bound of the loss

(♣) := $|\text{loss}(A(S); z) - \text{loss}(A(S'); z)|$ は次のように計算される:

$$\begin{aligned} (\clubsuit) &= \|\mathbf{b}(\mathbf{t}|P^{(K)}) - \mathbf{b}^*(\mathbf{t})\| - \|\mathbf{b}(\mathbf{t}|P'^{(K)}) - b^*(\mathbf{t})\| \\ &\stackrel{(1)}{=} \|\mathbf{b}(\mathbf{t}|P^{(K)} - P^*)\| - \|\mathbf{b}(\mathbf{t}|P'^{(K)} - P^*)\| \\ &\stackrel{(2)}{\leq} \|\mathbf{b}(\mathbf{t}|P^{(K)} - P^*) - \mathbf{b}(\mathbf{t}|P'^{(K)} - P^*)\| \\ &= \|(P^{(K)} - P^*)^\top z - (P'^{(K)} - P^*)^\top z\| \\ &= \|(P^{(K)} - P'^{(K)})^\top z\| \\ &\stackrel{(3)}{\leq} \|z\| \cdot \|P^{(K)} - P'^{(K)}\|. \end{aligned}$$

上から定数で押さえたい

- (1) Linearity of \mathbf{b}
- (2) Triangle inequality
- (3) Cauchy–Schwarz inequality

Derivation of the upper bound of $\|P^{(K)} - P'^{(K)}\|$

方針

- ① 条件数を用いて相対誤差を評価する
- ② 正則性を仮定して絶対誤差を評価する
- ③ PAC 学習における安定性による解析をする (未検討)

条件数と線型方程式の関係

線型方程式:

$$Ax = b, \quad (A + \Delta A)\tilde{x} = b + \Delta b$$

に対して、それぞれの解 x, \tilde{x} は以下の関係を満たす:

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \kappa(A) \frac{\|b - A\tilde{x}\|}{\|b\|}.$$

ここで、条件数 $\kappa(A)$ は、 $\kappa(A) := \lambda_{\max}(A)/\lambda_{\min}(A)$ で定義される。

相対誤差による評価

Proposition

$P^{(k+1)}, P'^{(k+1)}$ は次の線型方程式の解である.

$$\underbrace{(Z^{(k)})^\top Z^{(k)} P^{(k+1)}}_{\left((Z^{(k)})^\top Z^{(k)} + \Delta(Z^{(k)})^\top Z^{(k)}\right) \tilde{P}^{(k+1)}} = \underbrace{(Z^{(k)})^\top X^{(k+1)}}_{\left((Z^{(k)})^\top X^{(k+1)} + \Delta(Z^{(k)})^\top X^{(k+1)}\right)}.$$

ここで, Z, Z' および X, X' はそれぞれ一つの行のみ異なる行列である.

条件数と線型方程式の関係より, $P^{(k+1)}$ と $P'^{(k+1)}$ の相対誤差は

$$\frac{\|P^{(k+1)} - P'^{(k+1)}\|}{\|P^{(k+1)}\|} \leq \kappa \left((Z^{(k)})^\top Z^{(k)} \right) \frac{\|(Z^{(k)})^\top X^{(k+1)} - (Z^{(k)})^\top Z^{(k)} P'^{(k+1)}\|}{\|(Z^{(k)})^\top X^{(k+1)}\|}$$

のように評価できる.

絶対誤差による評価 I

Proposition

$P^{(k+1)}, P'^{(k+1)}$ は次の線型方程式の解である.

$$(Z^{(k)})^\top Z^{(k)} P^{(k+1)} = (Z^{(k)})^\top X^{(k+1)},$$
$$(Z'^{(k)})^\top Z'^{(k)} P'^{(k+1)} = (Z'^{(k)})^\top X'^{(k+1)}.$$

Proposition

正方形行列 $(Z^{(k)})^\top Z^{(k)}, (Z'^{(k)})^\top Z'^{(k)}$ の正則性を仮定すると,

$$P^{(k+1)} - P'^{(k+1)} = A^{-1}(Z^{(k)})^\top X^{(k+1)} - \tilde{A}^{-1}(Z'^{(k)})^\top X'^{(k+1)}$$

である. ここで, $A := (Z^{(k)})^\top Z^{(k)}$ であり, \tilde{A}^{-1} は以下で定義される.

$$\tilde{A}^{-1} := A^{-1} - A^{-1}U(I + VA^{-1}U)^{-1}VA^{-1}.$$

絶対誤差による評価 II

証明.

いま、行列 Z と Z' は一つの行のみ異なるような行列であるから、ある行列 U, V が存在して

$$(Z'^{(k)})^\top Z'^{(k)} = (Z^{(k)})^\top Z^{(k)} - UV$$

が成立する。ここで、Sherman-Morrison の公式¹を用いると

$$\begin{aligned} \left((Z'^{(k)})^\top Z'^{(k)} \right)^{-1} &= \left((Z^{(k)})^\top Z^{(k)} + UV \right)^{-1} \\ &= A^{-1} - A^{-1}U(I + VA^{-1}U)^{-1}VA^{-1} \end{aligned}$$

が成り立つ。ただし、 $A := (Z^{(k)})^\top Z^{(k)}$ とした。

¹ $(A + BDC)^{-1} = A^{-1} - A^{-1}BD(D + DCA^{-1}BD)^{-1}DCA^{-1}$

Derivation of the upper bound of $\|P^{(K)} - P'^{(K)}\|$

方針

- ① 条件数を用いて相対誤差を評価する
- ② 正則性を仮定して絶対誤差を評価する

- ① 相対誤差による評価:

$$\frac{\|P^{(k+1)} - P'^{(k+1)}\|}{\|P^{(k+1)}\|} \leq \kappa \left((Z^{(k)})^\top Z^{(k)} \right) \frac{\|(Z^{(k)})^\top X^{(k+1)} - (Z^{(k)})^\top Z^{(k)} P'^{(k+1)}\|}{\|(Z^{(k)})^\top X^{(k+1)}\|}.$$

- ② 絶対誤差による評価:

$$\|P^{(k+1)} - P'^{(k+1)}\| = \|A^{-1}(Z^{(k)})^\top X^{(k+1)} - \tilde{A}^{-1}(Z'^{(k)})^\top X'^{(k+1)}\|.$$

いずれも上から定数で押さえることができない・・・

Numerical experiments for $\|P^{(K)} - P'^{(K)}\|$

Numerical experiments

ノルム誤差 $\|P^{(K)} - P'^{(K)}\|$ の値の推移を実験的に観察する。

実験手順

- ① Δ^{M-1} から $\{t_i\}_{i=1}^{N+1}$ をサンプリングし, 2つの集合 $\mathcal{T}, \mathcal{T}'$ を定める:

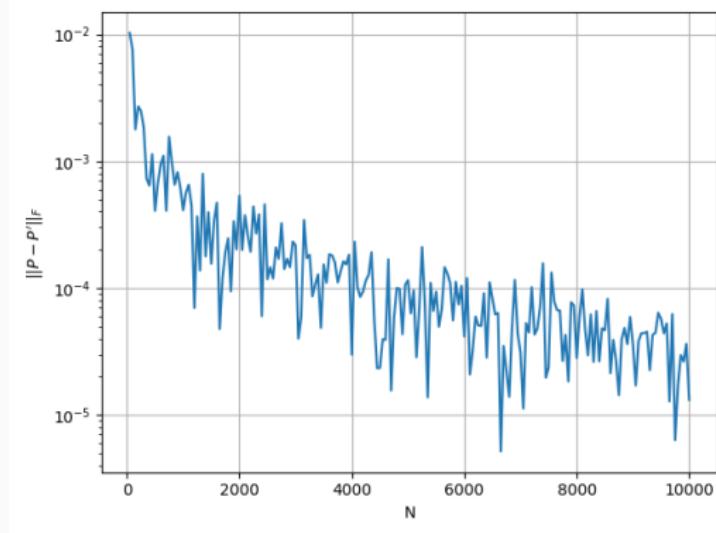
$$\mathcal{T} := \{ t_1, t_2, \dots, t_N \}, \quad \mathcal{T}' := \{ t_{N+1}, t_2, \dots, t_N \}.$$

- ② ベクトル z_i を計算し, 行列 Z, Z' を構成する
- ③ 提案手法で勾配降下した点集合をフィッティングするベジエ単体上から $\{x_i\}_{i=1}^{N+1}$ をサンプリングする
- ④ ベクトル x_i から行列 X, X' を構成し, 第一成分に対してガウシアンノイズ $\mathcal{N}(0, \sigma^2)$ を加える
- ⑤ 線型方程式を解き, P, P' を求める
- ⑥ フロベニウスノルム $\|P - P'\|_F$ を計算する

Numerical experiments for $\|P^{(K)} - P'^{(K)}\|$

Numerical experiments

ノルム誤差 $\|P^{(K)} - P'^{(K)}\|$ の値の推移を実験的に観察する。



- $N = 50k$ ($k = 1, 2, \dots, 200$) で実験
- $\mathcal{O}(1/N)$ で減少

Figure: $\|P^{(K)} - P'^{(K)}\|$ vs. sample size N