研究進捗報告書

ミーティング日： 2021 年 5 月 21 日　　学年 D2　　氏名 吉田　皓太郎

注意：ミーティング時には、必ず本報告書を作成し、一部を教員に提出すると共に、一部を自分用に持参して下さい。本報告書の提出がない場合、ミーティングは実施しません。また、項目１）から項目３）について未記入の箇所がある場合にも、ミーティングは実施しません。なお、本報告書は手書きでも構いません。

MTG を行うにあたってのねらい

今回の MTG では，局所修正を考慮したアルゴリズムを考慮し提案することを目標とする．これを踏まえて結果のディスカッションおよび，必要であれば修正・追加を行い，できれば論文投稿まで繋げたいと考えるため，その部分も併せてできればと思っている．

今週のミーティング事項について

目次

ミーティング事項の具体的な内容について

1　データ周りの話について

1.1　仮説

一般的には，十分にばらつきを持ったデータセットがあれば，回帰は誤差をほとんど持たずうまくいくはずである．今回のデータでは，パラメータのばらつきという部分は考慮しないまま学習させている．この事によって誤差が大きい所が発生していると考えられる．

つまり，今回のデータによる誤差は，データの疎密性の問題であると言い換えられるのではないだろうか．この仮説について，考察を行った．

1.2　前提

今回のパラメータは，関数 α, ω_η, D がそれぞれ，（仮定的に）ノイズを含んだ独立ガウス過程 $N(\mathbf{0}, \boldsymbol{K} + \sigma^2 \boldsymbol{I})$ に従って出力されているとする．$\boldsymbol{K}$ はグラム行列を表しており，各成分を定義するカーネル関数 $k(\boldsymbol{x}_i, \boldsymbol{x}_j)$ を設計することにより得る．

1.3　考察

まず初めに，特徴パラメータからガウス過程を用いて回帰させた際のハイパーパラメータに関する考察から行う．今回ガウス過程を適用する際に使用したカーネル関数は以下のように，RBF 関数と線形関数を組み合わせた ARD（関連度自動決定）カーネル関数である．ARD 関数は，パラメータ毎に重みが変化し，評価値に対する相関度を自動的に決定することができる．

$$k(\boldsymbol{x}_i, \boldsymbol{x}_j) = \theta_1 \exp\left[(\boldsymbol{x}_i - \boldsymbol{x}_j)^T diag(\boldsymbol{\Theta}_R)^{-1}(\boldsymbol{x}_i - \boldsymbol{x}_j)\right] + \boldsymbol{x}_i^T diag(\boldsymbol{\Theta}_L)\boldsymbol{x}_j \tag{1}$$

ここで $diag(\boldsymbol{\Theta})$ は，ベクトル Θ の各成分を対角成分に持つ対角行列である．ARD の場合は，$\dim \boldsymbol{\Theta}_L = \dim \boldsymbol{\Theta}_R = \dim \boldsymbol{x}$ となる．

また，今回特徴パラメータを抽出する際に用いたガウス過程のカーネル関数は RBF 関数であったことから，1 つの関数に

つき三つのパラメータ $\theta_1, \theta_2, \sigma$ を得る．今回得られた Θ_L, Θ_R を表にまとめると，次の通りとなった．

	α	ω_η	D
$\boldsymbol{\Theta}_{R,0}$	0.065274	0.047916	0.266596
$\boldsymbol{\Theta}_{R,1}$	158.2779	5.668934	0.286779
$\boldsymbol{\Theta}_{R,2}$	1	318.7759	1

	α	ω_η	D
$\boldsymbol{\Theta}_{L,0}$	9.827×10^{-3}	1.219×10^{-4}	1.083×10^{-6}
$\boldsymbol{\Theta}_{L,1}$	2.920×10^{1}	8.416×10^{-3}	1.420×10^{1}
$\boldsymbol{\Theta}_{L,2}$	1.000	2.732	1.000

この結果から，特に ω_η の分散に対して関連度が高いことが読み取れる．

データの疎密性がデータ間の距離に対して強く相関関係にある仮定すれば，疎密性を定義することは，データ間の距離をうまく定義することに等しくなるため，データ間の距離について考える．今回考察した距離はチェベシェフ距離，マンハッタン距離に加え，今回は α, ω_η, D をそれぞれ独立の確率過程により生成されるとすることから，確率間の距離である KL-Divergence を，各関数毎に考察を行った．

チェベシェフ距離，マンハッタン距離はそれぞれ次式で表される．

$$d_c(\boldsymbol{x}, \boldsymbol{y}) = \max_i \{|x_i - y_i|\} \tag{2}$$

$$d_m(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=0}^{n-1} |x_i - y_i| \tag{3}$$

また，KL-Divergence は次式で定義される．

$$d(p, q) = \int_{-\infty}^{\infty} p \log \left(\frac{p}{q} \right) dx \tag{4}$$

特に多変量正規分布の場合には

$$\frac{1}{2} \left[\log \frac{|\Sigma_1|}{|\Sigma_0|} - dim + Tr(\Sigma_1^{-1} \Sigma_0) + (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^T \Sigma_1^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) \right] \tag{5}$$

となる．

今回の数値検証では，検証データのパラメータ集合 $\boldsymbol{\Phi}_t$ と訓練データのパラメータ集合 $\boldsymbol{\Phi}_s$ において，各検証データのパラメータ $\boldsymbol{\phi}_{i,t} \in \boldsymbol{\Phi}_t$ に対して，データ間の距離 d_i を以下で定義した．

$$d_i = \min_j d_{method}(\boldsymbol{\phi}_{i,t}, \boldsymbol{\phi}_{j,s}) \tag{6}$$

これらのデータでそれぞれまとめたものを次の図で示す．また，先ほどの議論から，ω_η の分散と誤差の分布を示す． 各相関係数は以下になる．

	chebyshev	Manhattan	ω_η-Divergence	α-Divergence	D-Divergence	Relationship parameter
相関係数	0.665274	0.72512	0.266596	-0.053	0.0004218	0.474

2　次回の MTG について（終了後記載）

- ▼
- ▼

＃＃＃

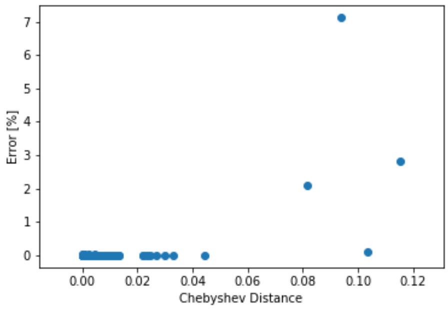

Fig. 1 Chebyshev

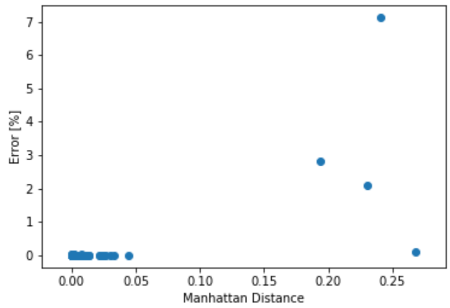

Fig. 2 Manhattan

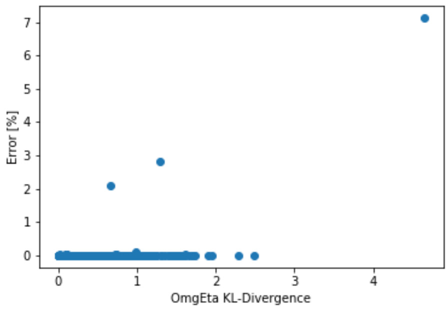

Fig. 3 ω_η-KL Divergence

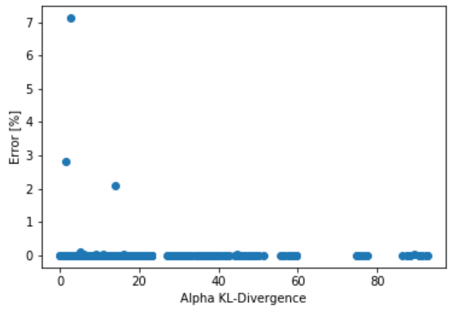

Fig. 4 α-KL Divergence

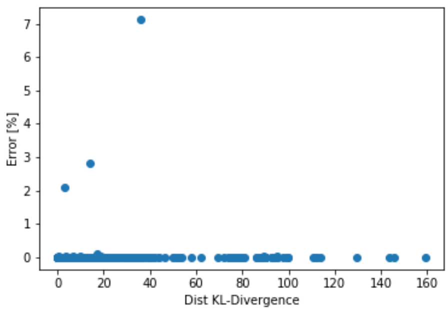

Fig. 5 D-KL Divergence

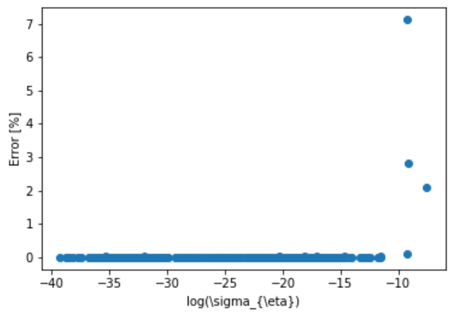

Fig. 6 関連度高いパラメータとの相関