数理解析チーム　太田 慎一

Mathematical Analysis Team

Shin-ichi Ohta

機械学習への幾何学的アプローチ

Contrastive Divergence法の漸近論的研究

藤平燎

$\mathcal{X}$ を有限集合，$\Theta\subset\mathbb{R}^d$ を開集合とする．
エネルギーベースモデル $\{p_\theta\}_{\theta\in\Theta}$：

$$p_\theta(x)=\frac{\exp(-H_\theta(x))}{Z(\theta)}\quad\left(Z(\theta)=\sum_{x\in\mathcal{X}}\exp(-H_\theta(x))\right)$$

と観測データ $X_1,\dots,X_N\sim p_{\theta_*}$ に対し，
最尤推定量を求める力学系は

$$\frac{\mathrm{d}\theta}{\mathrm{d}t}=\frac{\partial}{\partial\theta}\left(\frac{1}{N}\sum_{n=1}^{N}\log p_\theta(x)\right)=\left\langle\frac{\partial H_\theta}{\partial\theta}\right\rangle_{p_\theta}-\left\langle\frac{\partial H_\theta}{\partial\theta}\right\rangle_{\hat{p}_N}$$

データ数 $N\to\infty$ とすると，

$$\frac{\mathrm{d}\theta}{\mathrm{d}t}=\left\langle\frac{\partial H_\theta}{\partial\theta}\right\rangle_{p_\theta}-\left\langle\frac{\partial H_\theta}{\partial\theta}\right\rangle_{p_{\theta_*}}$$

このとき，θ_* は局所漸近安定な平衡点：
θ_* の近くから出発した軌道 $\theta(t)$ について，

$$\lim_{t\to\infty}\theta(t)=\theta_*$$

ただし，第1項は計算困難．

Contrastive Divergence法（Hinton, 2002）

詳細つりあい条件 $T_\theta(x\,|\,y)p_\theta(y)=T_\theta(y\,|\,x)p_\theta(x)$
をみたすマルコフ遷移カーネル $\{T_\theta\}_{\theta\in\Theta}$ に対して，

$$(\Phi_\theta\, q)(x):=\sum_{y\in\mathcal{X}}T_\theta(x\,|\,y)q(y)$$

とおき，

$$\frac{\mathrm{d}\theta}{\mathrm{d}t}=\left\langle\frac{\partial H_\theta}{\partial\theta}\right\rangle_{\Phi_\theta p_{\theta_*}}-\left\langle\frac{\partial H_\theta}{\partial\theta}\right\rangle_{p_{\theta_*}}\qquad(☆)$$

成果

θ_* は(☆)の局所漸近安定な平衡点．

結論

Contrastive Divergence法はデータ数 $N\to\infty$ のとき，
局所的には真のパラメータに収束する．

情報幾何との関係

一般に，(☆)の右辺は何らかの関数の勾配で表せない
→誘導される統計多様体構造に捩率が現れることと
関係している

Circle Patternと組合せリッチ流

離散データから連続データを復元する道具とそれを探すアルゴリズム

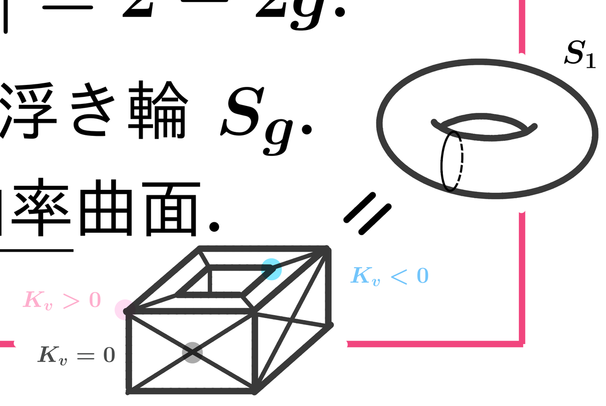
(A.Takatsu, to appear in Transactions of the American Mathematical Society)

離散閉曲面 T…3角形の繋ぎ合わせ．頂点間に相関有．

$\Theta:E\to[0,\frac{\pi}{2}]$

特徴.

- $\exists g\in\mathbb{N}$ s.t. $\chi(T):=\underset{\text{頂点集合}}{|V|}-\underset{\text{辺集合}}{|E|}+\underset{\text{面集合}}{|F|}=2-2g.$
- 位相的には，Tは滑らかな g個の穴あき浮き輪 S_g.
- 各3角形の大きさが適切⇝滑らかな定曲率曲面．
  不適⇝ ∃特異点．



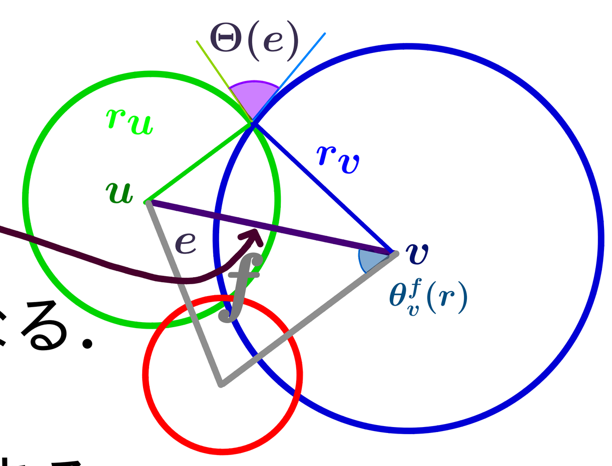
適切な大きさ＝Circle Pattern (Andreev’70, Thurston’78)

$r=(r_v)_{v\in V}\in\mathrm{R}^{V}_{>0}$ に対し

- $e\in E$ の長さが決まる．
- $f\in F$ は $\begin{Bmatrix}\mathbb{R}^2 & (g=1)\\ \mathbb{H}^2 & (g>1)\end{Bmatrix}$ の3角形になる．

  $\Rightarrow$ fの各頂点 v の角度 $\theta^f_v(r)$ が決まる．
- 組合せ曲率 $K_v(r):=2\pi-\sum_{v\in f\in F}\theta^f_v(r)$ が決まる．
- $K_v(r)=0\Rightarrow v$ での繋ぎは滑らか．

  $\therefore$ $K_v(r)=0,\ \forall v\in V\Rightarrow$滑らかな定曲率曲面をえる．



定義. $r\in\mathrm{R}^{V}{}_{>0}$ がCP :⇔ $K_v(r)=0,\ \forall v\in V$.

CPの探し方：組合せリッチ流 (Chow–Luo’03)

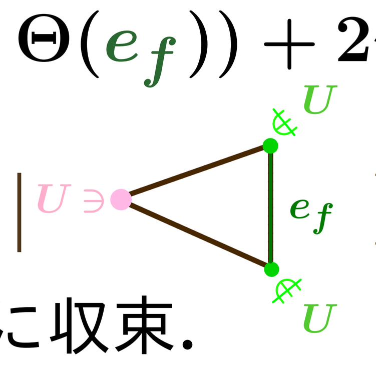
$$(*)\ \ \dot{r}(t)=-K(r(t))\sigma_g(r(t)),\quad \sigma_g(r):=\begin{cases}r & (g=1)\\ \sinh r & (g>1).\end{cases}$$

定理. (Thurston, Chow–Luo) (T,Θ) に対し以下は同値.:

- ∃!CP.
- $\emptyset\neq\forall U\subsetneq V$ に対し

  頂点集合がUのTの部分複体．

  $$\phi(U):=-\sum_{f\in\mathrm{Lk}(U)}(\pi-\Theta(e_f))+2\pi\chi(T_U)<0.$$

  $\mathrm{Lk}(U)=\{f\in F\ |$  $\}$

- $(*)$の解は$t\to\infty$で CP に収束．

問.(Chow–Luo) CPが存在しない(T,Θ)上の$(*)$の挙動は?

成果. (Takatsu)

仮定．(T,Θ)が $\min_{\emptyset\neq U\subsetneq V}\phi(U)=0$ を満たす．

結論．$(*)$の解$\{r(t)\}_{t\geq 0}$は

1. $r_v:=\lim_{t\to\infty}r_v(t)\in[0,\infty]$, $\lim_{t\to\infty}K_v(r(t))=0$ $(\forall v\in V)$.
2. ∃!離散曲面$T_0=(V_0,E_0,F_0)$とその相関Θ_0 s.t.
   $\chi(T_0)=\chi(T)$, $V_0\subset V$ かつ$(\tilde{r}_v)_{v\in V_0}$は(T_0,Θ_0)のCP.