異なるアルゴリズムを同じ視点から眺めてみよう

機械学習の幾何的解釈

NII

どんな研究？

様々なアルゴリズムを幾何的に捉えなおすことで，
今まで気が付かなかった知見を得る．

研究内容

変換 $\theta_{ij} = \log a_{ij} - \log a_{i-1,j} - \log a_{i,j-1} + \log a_{i-1,j-1}$

行列の世界

$A \in \mathbb{R}^{n \times n}$

各成分a_{ij}が行列を特徴づける

確率分布の世界

p_θ

パラメータθ_{ij}と
期待値η_{ij}で分布が決まる

$A_{ij} = \exp\left[\sum_{\substack{i' \le i \\ j' \le j}} \theta_{i'j'}\right]$

変換

Q.変換すると何が嬉しい？
A.行列に対する条件がθとηで簡単に記述できる！

例１ 行列の列和と行和を１に揃える（バランシング）

行列

Ask $\mathbf{u},\mathbf{v}$, *for*
$\mathrm{daig}(\mathbf{u}) A \mathrm{diag}(\mathbf{v}) \mathbf{1} = \mathbf{1}$
$\mathbf{1}_{ij} = 1$

確率

Sugiyama, M., et al
Tensor Balancing on Statistical Manifold,

$\eta_{1j} \leftarrow \frac{n-j}{n}$

例２ 任意の列を１列目の定数倍にする（1ランク近似）

行列

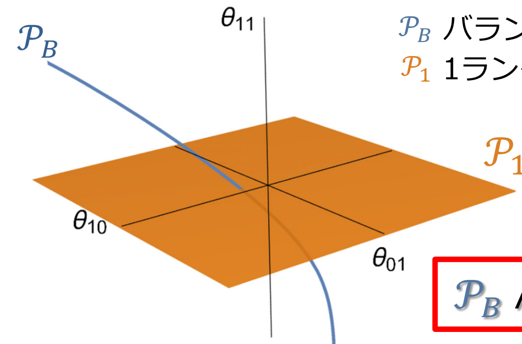
Ask A' *for*
$A' = \underset{\mathrm{rank}(A')=1}{\mathrm{argmin}} \; |A - A'|$

確率

$\theta_{ij} \leftarrow 0$
$(i,j) \in \{2,3,\cdots n\}^2$

分かったこと

(θ,η)の空間でアルゴリズムを議論して，
無関係に見える行列操作間の新しい視点を与える



$\mathcal{P}_B$ バランシング行列が居る曲線
$\mathcal{P}_1$ 1ランクの行列が居る曲面

$\mathcal{P}_B \wedge \mathcal{P}_1$ は一意に定まる.

(座標中の各点が行列に対応する)

他にも…

$\mathcal{P}_1$上では行列が分解できることが分かり，
物理学の平均場近似との類似性を指摘した．

できるようになったこと

θ,η座標系の凸な特徴を用いて
低ランク近似アルゴリズムの精度向上

20次元の行列の低ランク近似