識別的ベイズ離散ガウス過程潜在空間モデル

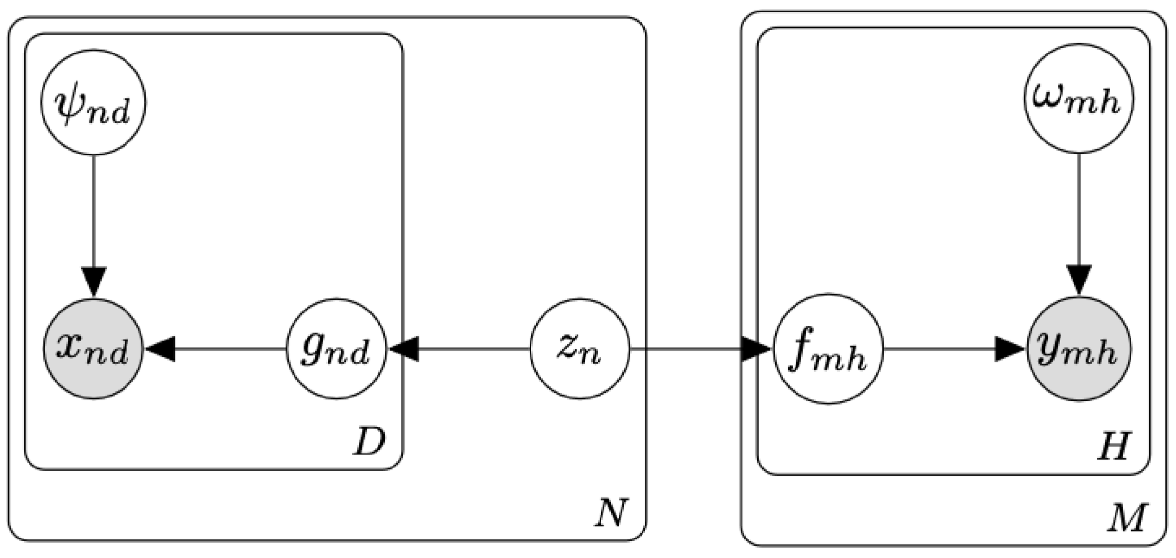
竹原 一彰　総合研究大学院大学 統計科学専攻

はじめに

推薦システムが提示するアイテムについてユーザーが受理するか否かを納得して判断をするには，なぜそれが推薦されたのかという解釈性が不可欠である．アイテムやユーザーの選好の分布，それらに対する推薦アイテムの位置づけなど全体像を捉えることが重要である．これを目的に潜在空間上での連続的なアイテム表現や選好分布を獲得できる識別的なベイズ離散ガウス過程潜在空間モデル（DDGPLVM）を提案する．

手法	次元削減	識別性	離散特徴量
PCA, t-SNE, GPLVM etc	O		
Urtasum et.al. [2]	O	O	
Gal et.al. [1]	O		O
本手法	O	O	O

識別的ベイズ離散ガウス過程潜在空間モデル



アイテム $X \in \{0,1\}^{N\times D}$，潜在空間上でのアイテム $Z \in \mathbb{R}^{N\times L}$，$H$ 人のユーザーの M 個のアイテムに対する二値評価 $Y \in \{0,1\}^{M\times H}$，$0 \le M \le N$，$G = \{g_{\cdot 1}, \dots, g_{\cdot D}\}, g_{\cdot d} \in \mathbb{R}^N$ と $F = \{f_{\cdot 1}, \dots, f_{\cdot H}\}, f_{\cdot h} \in \mathbb{R}^M$ は，それぞれガウス過程 $\mathcal{GP}(0, k_\phi)$，$\mathcal{GP}(0, k_\theta)$ に従う．$\Psi = (\psi_{nd}) \in \mathbb{R}^{N\times D}$ と $\Omega = (\omega_{mh}) \in \mathbb{R}^{M\times H}$ はPólya-Gamma 分布 $\mathcal{PG}$ に従う補助変数．

DDGPLVM によって得られた潜在空間 Z と選好 F の可視化を下図に示す．点=Z^{obs}，白抜き点=Z^{mis}，点サイズ=F（図(a)に関してはユーザー間の平均値）の大きさ，点の色=Y，例えば $H=3$ 人全員が高評価の場合 $(1,1,1)$ である．

推論アルゴリズム

Algorithm 1 DDGPLVM の推論アルゴリズム

Require: アイテム X，ユーザー評価 Y，潜在空間の次元 L，ガウス過程 F, G のカーネル関数 k_θ, k_ϕ，リープステップ数 N_{leap}，リープ単位 ϵ，イテレーション数 T，バーンイン期間 $S(<T)$

```
Initialize Z_0, Ω_0, Ψ_0
for t = 1, ..., T do
    Draw G_t ~ p(G|X, Ψ_{t-1}), F_t ~ p(F|Y, Ω_{t-1})
    Draw Ψ_t ~ p(Ψ|G_t), Ω_t ~ p(Ω|F_t)
    Draw Z_t^obs ~ p(Z^obs|F_t, G_t), Z_t^mis ~ p(Z^mis|G_t, Z_t^obs)
    Z_t ← Z_t^obs ∪ Z_t^mis
end for
return {Z_S, ..., Z_T}
```

$$p(\Omega|F) = \prod_{h=1}^{H}\prod_{m=1}^{M} \mathcal{PG}(\omega_{mh}|1, f_{mh}),\ p(\Psi|G) = \prod_{d=1}^{D}\prod_{n=1}^{N} \mathcal{PG}(\psi_{nd}|1, g_{nd})$$

$$p(F|Y,\Omega) \propto p(Y|F,\Psi)p(F)$$

$$\propto \prod_{h=1}^{H} \mathcal{N}(f_{\cdot h}|\eta_{\cdot h}^{\top}\Omega_h(\Omega_h + I_M)^{-1}, (\Omega_h + I_M)^{-1})$$

$$p(G|X,\Psi) \propto p(X|G,\Psi)p(G)$$

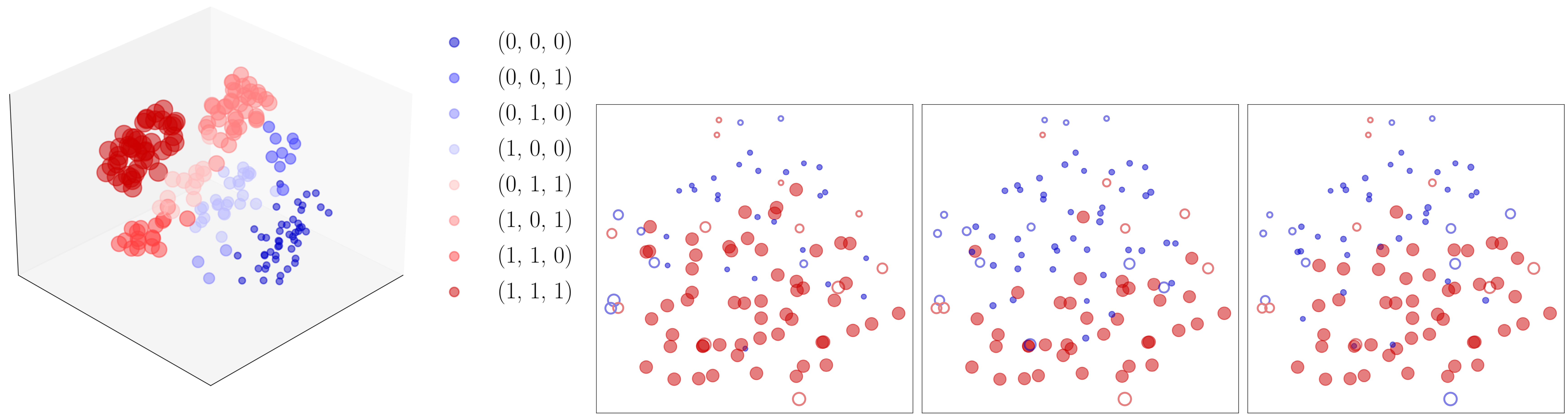
$$\propto \prod_{d=1}^{D} \mathcal{N}(g_{\cdot d}|\chi_{\cdot d}^{\top}\Psi_d(\Psi_d + I_N)^{-1}, (\Psi_d + I_N)^{-1})$$

$$p(Z^{obs}|F,G) \propto p(F|Z^{obs})p(G|Z^{obs})p(Z^{obs})$$

$$= \prod_{h=1}^{H} \mathcal{N}(f_{\cdot h}^{obs}|0, K_\theta^{obs}) \prod_{d=1}^{D} \mathcal{N}(g_{\cdot d}^{obs}|0, K_\phi^{obs}) \prod_{z_n \in Z^{obs}} \mathcal{N}(z_n|0, I_L)$$

$$p(Z^{mis}|G, Z^{obs}) \propto \int p(G^{mis}, G^{obs}|Z)dG^{obs}p(Z^{mis})$$

$$= \prod_{d=1}^{D} \int \mathcal{N}(g_{\cdot d}|0, K_\phi)dg_{\cdot d}^{obs} \prod_{z_n \in Z^{mis}} \mathcal{N}(z_n|0, I_L)$$



(a) 潜在空間 $Z(N = M = 200, H = 3)$　**(b)** ユーザー1$(N = 100, M = 80)$　**(c)** ユーザー2$(N = 100, M = 80)$　**(d)** ユーザー3$(N = 100, M = 80)$

参考文献

[1] Yarin Gal, Yutian Chen, and Zoubin Ghahramani. Latent gaussian processes for distribution estimation of multivariate categorical data. In *International Conference on Machine Learning*, pp. 645–654. jmlr.org, June 2015.

[2] Raquel Urtasun and Trevor Darrell. Discriminative gaussian process latent variable model for classification. In *Proceedings of the 24th international conference on Machine learning*, ICML '07, pp. 927–934, New York, NY, USA, June 2007. Association for Computing Machinery.