

§6.3 生成ネットワークの構造学習

- Deep learning モデルの学習

→ 性能の良いネットワークの構造を見つけるのが難しい.

→ **インド料理過程** (Indian buffet process) を使って,

データからネットワークの構造 (各層の幅や深さ, ユニットのつながりかた) を推定する.

6.3.1 インド料理過程

- **インド料理過程** (Indian buffet process):

列数が可算無限個のバイナリ行列 (要素が 0 か 1 の行列)

を生成するモデル.

6.3.1.1 無限行列の生成

- $M \in M_{N, H}(\{0, 1\})$ を考え, $H \rightarrow \infty$ のときの M の生成過程を構築.

[仮定] • M の第 n 列の要素は $m_{n,h} \stackrel{i.i.d.}{\sim} \text{Bern}(\pi_h)$ ($n=1, \dots, N$)

• 各 $h=1$ に対し, $\pi_h \stackrel{i.i.d.}{\sim} \text{Beta}(\frac{\alpha\beta}{H}, \beta)$ ($\alpha, \beta > 0, h=1, \dots, H$)

- M の分布は

$$\begin{aligned} p(M) &= \prod_{h=1}^H \prod_{n=1}^N p(m_{n,h}) \\ &= \prod_{h=1}^H \prod_{n=1}^N \int p(m_{n,h}, \pi_h) d\pi_h \end{aligned}$$

$$= \prod_{h=1}^H \int \prod_{n=1}^N p(m_{n,h}, \tau_n) d\tau_n$$

$$= \prod_{h=1}^H \int \prod_{n=1}^N p(m_{n,h} | \tau_n) p(\tau_n) d\tau_n$$

$$= \prod_{h=1}^H \int p(\tau_n) \prod_{n=1}^N p(m_{n,h} | \tau_n) d\tau_n$$

$$= \prod_{h=1}^H \int \frac{\Gamma(\frac{\alpha\beta}{H} + \beta)}{\Gamma(\frac{\alpha\beta}{H})\Gamma(\beta)} \tau_n^{\frac{\alpha\beta}{H}-1} (1-\tau_n)^{\beta-1} \prod_{n=1}^N \tau_n^{m_{n,h}} (1-\tau_n)^{1-m_{n,h}} d\tau_n.$$

$\therefore \tau_n$, $N_h := \sum_{n=1}^N m_{n,h}$ (第 h 列の 1 の個数) とすると,

$p(M)$

$$= \prod_{h=1}^H \frac{\Gamma(\frac{\alpha\beta}{H} + \beta)}{\Gamma(\frac{\alpha\beta}{H})\Gamma(\beta)} \int \tau_n^{N_h + \frac{\alpha\beta}{H} - 1} (1-\tau_n)^{N - N_h + \beta - 1} d\tau_n$$

\wedge β -分布の全範囲での積分が 1 になるのをこのように使う

$$= \prod_{h=1}^H \frac{\Gamma(\frac{\alpha\beta}{H} + \beta)}{\Gamma(\frac{\alpha\beta}{H})\Gamma(\beta)} \frac{\Gamma(N_h + \frac{\alpha\beta}{H})\Gamma(N - N_h + \beta)}{\Gamma(\frac{\alpha\beta}{H} + \beta + N)}$$

$$= \prod_{h=1}^H \frac{\Gamma(\frac{\alpha\beta}{H} + N_h)}{\Gamma(\frac{\alpha\beta}{H})} \cdot \frac{\Gamma(\frac{\alpha\beta}{H} + \beta)}{\Gamma(\frac{\alpha\beta}{H} + \beta + N)} \cdot \frac{\Gamma(N - N_h + \beta)}{\Gamma(\beta)} \quad \downarrow \Gamma(x+1) = x\Gamma(x)$$

$$= \prod_{h=1}^H \left(\prod_{n=0}^{N_h-1} \left(\frac{\alpha\beta}{H} + n \right) \right) \left(\prod_{n=0}^{N-1} \frac{1}{\frac{\alpha\beta}{H} + \beta + n} \right) \left(\prod_{n=0}^{N-N_h-1} (\beta + n) \right)$$

$\downarrow \Rightarrow$ くりこみ

$$= \prod_{h=1}^H \left(\prod_{n=0}^{N_h-1} \left(\frac{\alpha\beta}{H} + n \right) \right) \left(\prod_{n=0}^{N_h-1} \frac{1}{\frac{\alpha\beta}{H} + \beta + n} \right) \left(\prod_{n=0}^{N-N_h-1} \frac{1}{\frac{\alpha\beta}{H} + \beta + N_h + n} \right) \left(\prod_{n=0}^{N-N_h-1} (\beta + n) \right)$$

$$= \prod_{h=1}^H \left(\prod_{n=0}^{N_h-1} \frac{\frac{\alpha\beta}{H} + n}{\frac{\alpha\beta}{H} + n + \beta} \right) \left(\prod_{n=0}^{N-N_h-1} \frac{\beta + n}{\frac{\alpha\beta}{H} + N_h + \beta + n} \right)$$

$\in (0,1)$

$\in (0,1)$

$$\therefore \prod_{h=1}^H \frac{K_h}{\in (0,1)} \xrightarrow{H \rightarrow \infty} 0.$$

$H \rightarrow \infty$ で全てのバイナリ行列の生成確率が 0 になってしまう.

これを防ぐため,

$$[M] := \{ M' \in M_{N,H}(\{0,1\}) \mid M' \text{ の列を並びかざると } M \text{ と一致する} \}$$

と行列の同値類をつくり, $[M]$ に関する分布を考えることにする.

cf.) 同値類.

集合 X における二項関係 \sim に対し $\forall \alpha, \gamma, \delta \in X$ に対し

$$(1) \alpha \sim \alpha \quad (2) \alpha \sim \gamma \Rightarrow \gamma \sim \alpha \quad (3) \alpha \sim \gamma, \gamma \sim \delta \Rightarrow \alpha \sim \delta$$

をみたすとき, \sim を同値関係という. (e.g., $=$, 図形の相似関係 など)

集合 $[\alpha] := \{ \alpha' \in X \mid \alpha' \sim \alpha \}$ を α の同値類という.

$\alpha' \in [\alpha]$ を一つ選ぶとき, α' を同値類 $[\alpha]$ の代表元という.

同値類を全てあつめた集合 $X/\sim := \{ [\alpha] \mid \alpha \in X \}$ を \sim による X の商集合という.

特に, $[M]$ の代表元 M' を次のとおりに選ぶ:

M の各列を N 桁の 2 進数とみて, 数値の大きい順に左の列から並べ

ておく行列を M' とする.

left-ordered form
 $M' = \text{lof}(M)$ と書く.

M の列を N 桁の 2 進数とみて (10 進数で) i になるものの個数を

H_i ($i = 0, 1, \dots, 2^N - 1$) と書く.

同じものを含む順列の数を表す。

- 同値類 $[M]$ に含まれる行列の個数は、多項係数を考え

$$|[M]| = \binom{H}{H_0, H_1, \dots, H_{2^M-1}} = \frac{H!}{\prod_{i=0}^{2^M-1} H_i!} \quad \square.$$

- 以上より,

$$\begin{aligned} p([M]) &= \sum_{M' \in [M]} p(M) \\ &= \frac{H!}{\prod_{i=0}^{2^M-1} H_i!} \prod_{h=1}^H \frac{\Gamma(\frac{\alpha\beta}{H} + \beta)}{\Gamma(\frac{\alpha\beta}{H})\Gamma(\beta)} \frac{\Gamma(N_h + \frac{\alpha\beta}{H})\Gamma(N - N_h + \beta)}{\Gamma(\frac{\alpha\beta}{H} + \beta + N)}. \end{aligned}$$

- ここで,

$$H_+ := H - H_0 \quad (N_h > 0 \text{ とする列の個数})$$

とすると, (left-ordered form に並べたものを考え)

$$p([M])$$

$$= \frac{H!}{\prod_{i=0}^{2^M-1} H_i!} \left(\frac{\Gamma(N+\beta)\Gamma(\frac{\alpha\beta}{H}+\beta)}{\Gamma(\beta)\Gamma(\frac{\alpha\beta}{H}+\beta+N)} \right)^{H_0} \prod_{h=1}^{H_+} \frac{\Gamma(\frac{\alpha\beta}{H}+\beta)}{\Gamma(\frac{\alpha\beta}{H})\Gamma(\beta)} \frac{\Gamma(N_h + \frac{\alpha\beta}{H})\Gamma(N - N_h + \beta)}{\Gamma(\frac{\alpha\beta}{H} + \beta + N)}$$

$N_h > 0$ とする h の数.

$$= \frac{H!}{\prod_{i=0}^{2^M-1} H_i!} \left(\frac{\Gamma(N+\beta)\Gamma(\frac{\alpha\beta}{H}+\beta)}{\Gamma(\beta)\Gamma(\frac{\alpha\beta}{H}+\beta+N)} \right)^{H_0} \left(\frac{\Gamma(N+\beta)\Gamma(\frac{\alpha\beta}{H}+\beta)}{\Gamma(\beta)\Gamma(\frac{\alpha\beta}{H}+\beta+N)} \right)^{H_+} \prod_{h=1}^{H_+} \frac{\Gamma(N_h + \frac{\alpha\beta}{H})\Gamma(N - N_h + \beta)}{\Gamma(\frac{\alpha\beta}{H})\Gamma(N+\beta)}$$

$$= \frac{H!}{\prod_{i=0}^{2^M-1} H_i!} \left(\prod_{n=0}^{N-1} \frac{\beta+n}{\frac{\alpha\beta}{H} + \beta + n} \right)^H \prod_{h=1}^{H_+} \left(\prod_{n=0}^{N_h-1} \frac{(\frac{\alpha\beta}{H} + n)}{\Gamma(N+\beta)} \frac{\Gamma(N - N_h + \beta)}{\Gamma(N+\beta)} \right)$$

$\Gamma(x+1) = x\Gamma(x)$

$$= \frac{H!}{\prod_{i=0}^{2^M-1} H_i!} \left(\prod_{n=0}^{N-1} \frac{\beta+n}{\frac{\alpha\beta}{H} + \beta + n} \right)^H \prod_{h=1}^{H_+} \left(\frac{\alpha\beta}{H} \prod_{n=1}^{N_h-1} \left(\frac{\alpha\beta}{H} + n \right) \frac{\Gamma(N - N_h + \beta)}{\Gamma(N+\beta)} \right)$$

$$= \frac{H!}{\prod_{i=0}^{2^M-1} H_i!} \left(\prod_{n=0}^{N-1} \frac{\beta+n}{\frac{\alpha\beta}{H} + \beta + n} \right)^H \left(\frac{\alpha\beta}{H} \right)^{H_+} \prod_{h=1}^{H_+} \left(\prod_{n=1}^{N_h-1} \left(\frac{\alpha\beta}{H} + n \right) \frac{\Gamma(N - N_h + \beta)}{\Gamma(N+\beta)} \right)$$

$$= \frac{(\alpha\beta)^{H_+}}{\prod_{i=1}^{2^p-1} H_i!} \frac{H!}{H_0! H^{H_+}} \left(\prod_{n=0}^{N-1} \frac{\beta+n}{\frac{\alpha\beta}{H} + \beta+n} \right)^H \prod_{h=1}^{H_+} \left(\prod_{n=1}^{N_h-1} \left(\frac{\alpha\beta}{H} + n \right) \frac{\Gamma(N-N_h+\beta)}{\Gamma(N+\beta)} \right).$$

よって,

$$\frac{H!}{H_0! H^{H_+}} = \frac{1}{H^{H_+}} \prod_{h=1}^{H_+} (H-h+1) = \prod_{h=1}^{H_+} \left(1 - \frac{h}{H} + \frac{1}{H} \right) \xrightarrow{H \rightarrow \infty} 1.$$

また,

$$\begin{aligned} \left(\prod_{n=0}^{N-1} \frac{\beta+n}{\frac{\alpha\beta}{H} + \beta+n} \right)^H &= \left(\prod_{n=0}^{N-1} \left(\frac{\alpha\beta}{H(\beta+n)} + 1 \right)^{-1} \right)^H = \prod_{n=0}^{N-1} \left(1 + \frac{\alpha\beta}{H(\beta+n)} \right)^{-H} \\ \xrightarrow{H \rightarrow \infty} \prod_{n=0}^{N-1} \exp\left(-\frac{\alpha\beta}{\beta+n}\right) &= \exp\left(-\alpha \sum_{n=0}^{N-1} \frac{\beta}{n+\beta}\right) = \exp\left(-\alpha \sum_{n=1}^N \frac{\beta}{n+\beta-1}\right) \\ &= \exp(-\bar{H}_+). \end{aligned}$$

よって,

$$\prod_{n=1}^{N_h-1} \left(\frac{\alpha\beta}{H} + n \right) \xrightarrow{H \rightarrow \infty} (N_h-1)! = \Gamma(N_h).$$

よって,

$$p([M]) \xrightarrow{H \rightarrow \infty} \frac{(\alpha\beta)^{H_+}}{\prod_{i=1}^{2^p-1} H_i!} \exp(-\bar{H}_+) \prod_{h=1}^{H_+} \frac{\Gamma(N_h) \Gamma(N-N_h+\beta)}{\Gamma(N+\beta)} =: p_\infty([M]).$$

この分布は M の行を交換しても ($\prod_{i=1}^{2^p-1} H_i!$, H_+ , N_h が変わらないので)

確率が変わらない (交換可能性をもつ)。

Remark \bar{H}_+ は H_+ の期待値: $\bar{H}_+ = \mathbb{E}_{p_\infty}[H_+]$.

pf. $\nu_i :=$ (この2進数表示における1の個数) とする。

$p_\infty([M])$

$$= \frac{(\alpha\beta)^{H_+}}{\prod_{i=1}^{2^N-1} H_i!} \exp(-\bar{H}_+) \prod_{h=1}^{H_+} \frac{\Gamma(N_h)\Gamma(N-N_h+\beta)}{\Gamma(N+\beta)}$$

第h列 (h=1, ..., H_+) は
2進数で i (i=1, ..., 2^N-1) を表す. i を表す列は H_i コ.
第h列の i を表すなら, N_h = \nu_i.

$$= \frac{(\alpha\beta)^{\sum_{i=1}^{2^N-1} H_i}}{\prod_{i=1}^{2^N-1} H_i!} \exp(-\bar{H}_+) \prod_{i=1}^{2^N-1} \left(\frac{\Gamma(\nu_i)\Gamma(N-\nu_i+\beta)}{\Gamma(N+\beta)} \right)^{H_i}$$

$$= \exp(-\bar{H}_+) \prod_{i=1}^{2^N-1} \frac{1}{H_i!} \left(\alpha\beta \frac{\Gamma(\nu_i)\Gamma(N-\nu_i+\beta)}{\Gamma(N+\beta)} \right)^{H_i}$$

\(\Gamma(x+1) = x\Gamma(x)\)

$$= \exp(-\bar{H}_+) \prod_{i=1}^{2^N-1} \frac{1}{H_i!} \left(\alpha\beta \frac{(\nu_i-1)!}{\prod_{n=1}^{\nu_i} (N-n+\beta)} \right)^{H_i}$$

=: \(\lambda_i\)

$$= \exp(-\bar{H}_+ + \sum_{i=1}^{2^N-1} \lambda_i) \prod_{i=1}^{2^N-1} \frac{\lambda_i^{H_i} e^{-\lambda_i}}{H_i!}$$

$$= \exp(-\bar{H}_+ + \sum_{i=1}^{2^N-1} \lambda_i) \prod_{i=1}^{2^N-1} \text{Poi}(H_i | \lambda_i)$$

= const.

各 H_i は独立にパラメータ λ_i の Poisson 分布に従っている。

$p_\infty([M])$ を各 H_i に対する同時分布とみると, 係数 $\exp(-\bar{H}_+ + \sum_{i=1}^{2^N-1} \lambda_i)$

は 1 でないといわない。つまり,

$$\begin{aligned} \bar{H}_+ &= \sum_{i=1}^{2^N-1} \lambda_i \\ &= \sum_{i=1}^{2^N-1} \mathbb{E}_{\text{Poi}(H_i|\lambda_i)}[H_i] \\ &= \sum_{i=1}^{2^N-1} \mathbb{E}_{p_\infty}[H_i] \\ &= \mathbb{E}_{p_\infty} \left[\sum_{i=1}^{2^N-1} H_i \right] \\ &= \mathbb{E}_{p_\infty}[H_+] \end{aligned}$$



・ 実際には、 $p_\infty([M])$ に従うような行列 M をサンプリングするには、次の

インド料理過程 IBP(α, β) に従ってサンプリングすればよい:

$$N_h^{(n)} := \sum_{i=1}^n m_{i,h} \quad (n=1, \dots, N) \text{ とする.}$$

1. $\eta_1 \sim \text{Poi}(\alpha)$; $m_{1,h} \leftarrow 1 \quad (h=1, \dots, \eta_1)$; $H \leftarrow \eta_1$. ← 客1ははじめから η_1 コの料理をとり.

2. 各 $n \quad (n=2, 3, \dots, N)$ に対し、次を行: ← 客 n の挙動

$$m_{n,h} \sim \text{Bern}\left(\frac{N_h^{(n-1)}}{n+\beta-1}\right) \quad (h=1, \dots, H);$$

各料理 $h=1, \dots, H$ をとる確率は、
← 既に1と付いているものが多いほど1に近づいたり、
(人気のある料理ほど選ばれる)

$$\eta_n \sim \text{Poi}\left(\frac{\alpha\beta}{n+\beta-1}\right);$$

$$m_{n,H+h} \leftarrow 1 \quad (h=1, \dots, \eta_n);$$

← 新しい料理を η_n コとる.

$$H \leftarrow H + \eta_n;$$

・ これで生成される M の同値類 $[M]$ の生成確率は $p_\infty([M])$ になる.

pf. N についての帰納法. N のときの $p_\infty([M]) \in p_\infty^{(N)}([M])$, $H_+ \in H_+^{(N)}$ と書き,

M の n 行目までの行列を $M^{(n)}$ と書く ($n=1, \dots, N$)

・ $N=1$ のとき, $H_+^{(1)} = H_1$, $\bar{H}_+ = \alpha + 2\alpha z$.

$$p_\infty^{(1)}([M]) = \frac{(\alpha\beta)^{H_1}}{H_1!} \exp(-\alpha) \prod_{h=1}^{H_1} \frac{\Gamma(1)\Gamma(\beta)}{\Gamma(1+\beta)} = \frac{\alpha^{H_1}}{H_1!} e^{-\alpha} = \text{Poi}(\alpha)$$

= $\frac{\alpha}{\beta}$

・ N のときの成立を仮定.

$N+1$ のとき, 仮定より $[M^{(N)}]$ の生成確率は $p_\infty^{(N)}([M^{(N)}])$ である.

$$p_\infty^{(N+1)}([M]) = \frac{(\alpha\beta)^{H_+^{(N+1)}}}{\prod_{i=1}^{N+1} H_i!} \exp\left(-\alpha \sum_{n=1}^N \frac{\beta}{n+\beta-1} - \alpha \frac{\beta}{N+\beta}\right) \prod_{h=1}^{H_+^{(N+1)}} \frac{\Gamma(N_h) \Gamma(N+1-N_h+\beta)}{\Gamma(N+1+\beta)}$$

$$\begin{aligned}
&= \left(\prod_{i=1}^{2^N-1} \frac{(\alpha\beta)^{H_i}}{H_i!} \right) \left(\prod_{i=2^N}^{2^{M+1}-1} \frac{(\alpha\beta)^{H_i}}{H_i!} \right) \exp\left(-\alpha \sum_{n=1}^N \frac{\beta}{n+\beta-1}\right) \exp\left(-\alpha \frac{\beta}{N+\beta}\right) \\
&\quad \times \prod_{i=1}^{2^N-1} \left(\frac{\Gamma(\nu_i)\Gamma(N-\nu_i+\beta)}{\Gamma(N+\beta)} \right)^{H_i} \prod_{i=1}^{2^N-1} \left(\frac{N-\nu_i+\beta}{N+\beta} \right)^{H_i} \prod_{i=2^N}^{2^{M+1}-1} \left(\frac{\Gamma(\nu_i)\Gamma(N+1-\nu_i+\beta)}{\Gamma(N+1+\beta)} \right)^{H_i} \\
&= P_{\infty}^{(N)}([M^{(N)}]) \cdot \frac{(\alpha\beta)^{H_{2^N}}}{H_{2^N}!} \exp\left(-\alpha \frac{\beta}{N+\beta}\right) \left(\frac{\Gamma(1)\Gamma(N+\beta)}{\Gamma(N+1+\beta)} \right)^{H_{2^N}} \\
&\quad \times \left(\prod_{i=1}^{2^N-1} \left(\frac{N-\nu_i+\beta}{N+\beta} \right)^{H_i} \right) \left(\prod_{i=1}^{2^N-1} \frac{1}{H_{2^{N+i}}!} \left(\alpha\beta \frac{\Gamma(\nu_{2^{N+i}})\Gamma(N+1-\nu_{2^{N+i}}+\beta)}{\Gamma(N+1+\beta)} \right)^{H_{2^{N+i}}} \right) \\
&= P_{\infty}^{(N)}([M^{(N)}]) \cdot P_{0i}(H_{2^N} | \frac{\alpha\beta}{N+\beta}) \prod_{i=1}^{2^N-1} \left(\left(1 - \frac{\nu_i}{N+\beta}\right)^{H_i} \frac{1}{H_{2^{N+i}}!} \left(\alpha\beta \prod_{n=0}^{\nu_i-1} \frac{\nu_i-n}{N-n+\beta} \right)^{H_{2^{N+i}}} \right) \\
&= P_{\infty}^{(N)}([M^{(N)}]) \cdot P_{0i}(H_{2^N} | \frac{\alpha\beta}{N+\beta}) \prod_{i=1}^{2^N-1} \left(\left(\frac{\nu_i}{N+\beta} \right)^{H_{2^{N+i}}} \left(1 - \frac{\nu_i}{N+\beta}\right)^{H_i} \right) \\
&\quad \times \prod_{i=1}^{2^N-1} \frac{1}{H_{2^{N+i}}!} \left(\alpha\beta (\nu_i-1)! \prod_{n=1}^{\nu_i-1} \frac{1}{N-n+\beta} \right)^{H_{2^{N+i}}}
\end{aligned}$$

$\nu_{2^{N+i}} = \nu_i + 1$. ($\nu_i > 0$)

... 72 と 74 は 73 を (72) を示せばいい。

• IBPで生成されるバイナリ行列 M の性質

1. 各行の 1 の個数は $Poi(\alpha)$ に従う.

pf. 1行目の 1 の個数は $Poi(\alpha)$ に従う. 交換可能性からあるので,

他の行もそうである. \square

2. 1 の要素の個数の期待値は $\sum_{h=1}^{\infty} E_{p_{\alpha}}[N_h] = N\alpha$ Bern(α) の期待値.

$$\begin{aligned} \text{pf. } \sum_{h=1}^{\infty} E[N_h] &= \lim_{H \rightarrow \infty} \sum_{h=1}^H E_{p^{(H)}}[N_h] = \lim_{H \rightarrow \infty} \sum_{h=1}^H \sum_{n=1}^N E_{p^{(H)}}[m_{n,h}] \\ &= \lim_{H \rightarrow \infty} \sum_{h=1}^H \sum_{n=1}^N \frac{\frac{\alpha \beta}{H}}{\frac{\alpha \beta}{H} + \beta} \\ &= \lim_{H \rightarrow \infty} \frac{N\alpha}{\frac{\alpha}{H} + 1} \\ &= N\alpha. \quad \square \end{aligned}$$

3. $E_{p_{\alpha}}[H_+] = \bar{H}_+$. (既示 = 示した)

4. $\lim_{\beta \rightarrow 0} \bar{H}_+ = \alpha$

$$\begin{aligned} \text{pf. } \lim_{\beta \rightarrow 0} \bar{H}_+ &= \lim_{\beta \rightarrow 0} \alpha \sum_{n=1}^N \frac{\beta}{n + \beta - 1} = \lim_{\beta \rightarrow 0} \alpha \sum_{n=1}^N \frac{1}{1 + \frac{n-1}{\beta}} \\ &= \lim_{\beta \rightarrow 0} \left(\alpha + \alpha \sum_{n=2}^N \frac{1}{1 + \frac{n-1}{\beta}} \right) = \alpha. \quad \square \end{aligned}$$

5. $\lim_{\beta \rightarrow \infty} \bar{H}_+ = N\alpha$

$$\text{pf. } \lim_{\beta \rightarrow \infty} \bar{H}_+ = \lim_{\beta \rightarrow \infty} \left(\alpha + \alpha \sum_{n=2}^N \frac{1}{1 + \frac{n-1}{\beta}} \right) = N\alpha. \quad \square$$