機械学習 Machine Learning

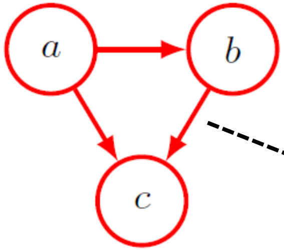
グラフィカルモデル Graphical Models

福嶋 誠 Makoto Fukushima

情報科学部
School of Informatics and Data Science

グラフィカルモデル Graphical models
有向グラフ Directed graphs

グラフィカルモデルによって確率分布の構造を図示することができる.
Graphical models allows structured probability distributions to be expressed in graphical form.

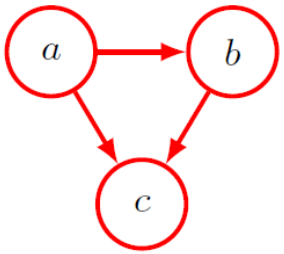


ノード：確率変数
Nodes: random variables

リンク：変数間の確率的な関係
Links: probabilistic relationships between variables

グラフのリンクが矢印による特定の向きをもつ有向グラフィカルモデルに焦点をあてる.
We will focus on directed graphical models in which the links of the graphs have a particular direction indicated by arrows.

グラフィカルモデル Graphical models
分解 Factorization



確率の乗法定理を適用
Application of the product rule of probability

$$p(a, b, c) = p(c|a, b)p(a, b) = p(c|a, b)p(b|a)p(a)$$

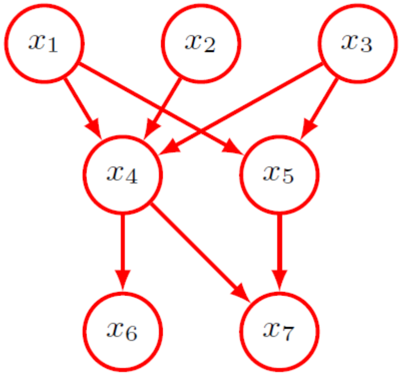
ノードa はノードb の親.
Node a is the parent of node b.

ノードb はノードa の子.
Node b is the child of node a.

K 個の変数の同時分布に対しては,
For the joint distribution over K variables,

$$p(x_1, \ldots, x_K) = p(x_K|x_1, \ldots, x_{K-1}) \ldots p(x_2|x_1)p(x_1)$$

グラフィカルモデル Graphical models
分解 Factorization



グラフが全結合していないケース
In a case where the graph is not fully connected

$$
\begin{aligned}
p(x_1, \ldots, x_7) \quad = \quad & p(x_1)p(x_2)p(x_3)p(x_4|x_1, x_2, x_3) \\
& p(x_5|x_1, x_3)p(x_6|x_4)p(x_7|x_4, x_5)
\end{aligned}
$$

一般的な分解
General factorization

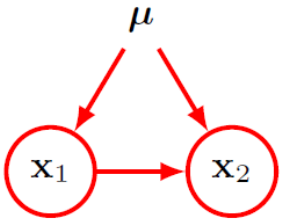
$$p(x_1, \ldots, x_K) = \prod_{k=1}^{K} p(x_k|\mathrm{pa}(k))$$

$\mathrm{pa}(k)$: x_k の親の集合 The set of parents of x_k

ここで考えているグラフは有向閉路をもたないという制約を満たす（有向非循環グラフ）.
In the graph we are considering, there are no closed path (directed acyclic graph).

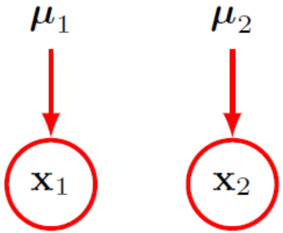
グラフィカルモデル Graphical models
離散変数　Discrete variables

一般的な同時分布　K^2-1 個のパラメータ
General joint distribution　K^2-1 parameters



$$p(\mathbf{x}_1, \mathbf{x}_2|\boldsymbol{\mu}) = \prod_{k=1}^{K}\prod_{l=1}^{K} \mu_{kl}^{x_{1k}x_{2l}}$$

独立な同時分布　$2(K-1)$ 個のパラメータ
Independent joint distribution　$2(K-1)$ parameters



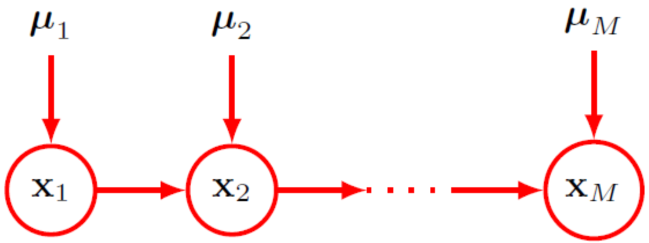
$$\hat{p}(\mathbf{x}_1, \mathbf{x}_2|\boldsymbol{\mu}) = \prod_{k=1}^{K} \mu_{1k}^{x_{1k}} \prod_{l=1}^{K} \mu_{2l}^{x_{2l}}$$

グラフィカルモデル Graphical models
離散変数 Discrete variables

一般的な同時分布（変数がM個ある場合） $K^M - 1$ 個のパラメータ
General joint distribution over M variables $K^M - 1$ parameters
（グラフが全結合している場合 If the graph is fully connected）

独立な同時分布（変数がM個ある場合） $M(K-1)$ 個のパラメータ
Independent joint distribution over M variables $M(K-1)$ parameters

以下のM個のノードの連鎖の場合
The following chain of M nodes



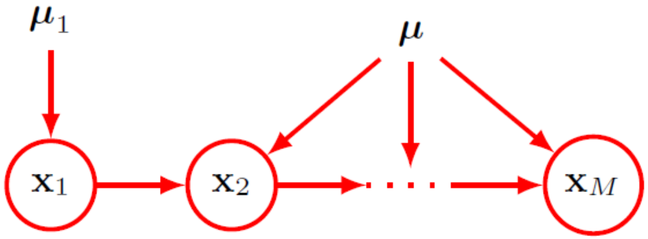
$K-1+(M-1)K(K-1)$ 個のパラメータ
$K-1+(M-1)K(K-1)$ parameters

グラフィカルモデル Graphical models
離散変数　Discrete variables

一般的な同時分布（変数がM 個ある場合）　$K^M - 1$ 個のパラメータ
General joint distribution over M variables　$K^M - 1$ parameters
（グラフが全結合している場合 If the graph is fully connected）

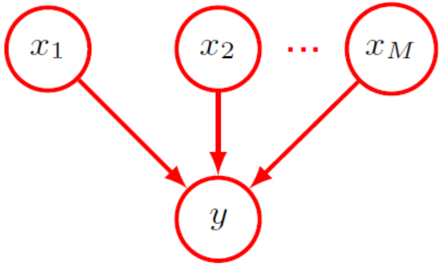
独立な同時分布（変数がM 個ある場合）　$M(K-1)$ 個のパラメータ
Independent joint distribution over M variables　$M(K-1)$ parameters

以下のM 個のノードの連鎖の場合
The following chain of M nodes



$K - 1 + K(K-1) = K^2 - 1$ 個のパラメータ
$K - 1 + K(K-1) = K^2 - 1$ parameters

グラフィカルモデル Graphical models
離散変数　Discrete variables



$x_1, \ldots, x_M$ がK 個の状態を取り得る離散変数のとき，
$p(y=1|x_1, \ldots, x_M)$ は一般的に$O(K^M)$ 個のパラメータをもつ．

If $x_1, \ldots, x_M$ are discrete K-state variables,
$p(y=1|x_1, \ldots, x_M)$ in general has $O(K^M)$ parameters.

以下でパラメタライズされた形式の場合は，
The parametrized form

$$p(y=1|x_1, \ldots, x_M) = \sigma\left(w_0 + \sum_{i=1}^{M} w_i x_i\right) = \sigma(\mathbf{w}^{\mathrm{T}}\mathbf{x})$$

$M+1$ 個のパラメータしか必要としない．
requires only $M+1$ parameters.

グラフィカルモデル Graphical models
ガウス変数 Gaussian variables

$$p(x_i|\mathrm{pa}_i) = \mathcal{N}\left(x_i \middle| \sum_{j\in\mathrm{pa}_i} w_{ij}x_j + b_i, v_i\right)$$

各ノードはガウス変数であり，
その平均は親ノードの線形関数で表される．
Each node is Gaussian and the mean
is a linear function of the parents.

同時分布の対数 The log of the joint distribution

$$\ln p(\mathbf{x}) = \sum_{i=1}^{D} \ln p(x_i|\mathrm{pa}(i)) = -\sum_{i=1}^{D} \frac{1}{2v_i}\left(x_i - \sum_{j\in\mathrm{pa}(i)} w_{ij}x_j - b_i\right)^2 + \mathrm{const}$$

平均と共分散 Mean and covariance

$$\mathbb{E}[x_i] = \sum_{j\in\mathrm{pa}(i)} w_{ij}\mathbb{E}[x_j] + b_i \qquad \mathrm{cov}[x_i, x_j] = \sum_{k\in\mathrm{pa}(j)} w_{jk}\mathrm{cov}[x_i, x_k] + I_{ij}v_j$$

グラフィカルモデル Graphical models
ガウス変数　Gaussian variables

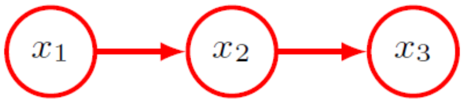
$$p(x_i|\mathrm{pa}_i) = \mathcal{N}\left(x_i \,\middle|\, \sum_{j\in\mathrm{pa}_i} w_{ij}x_j + b_i, v_i\right)$$

各ノードはガウス変数であり，
その平均は親ノードの線形関数で表される．
Each node is Gaussian and the mean
is a linear function of the parents.

平均と共分散 Mean and covariance

$$\mathbb{E}[x_i] = \sum_{j\in\mathrm{pa}(i)} w_{ij}\mathbb{E}[x_j] + b_i \qquad \mathrm{cov}[x_i, x_j] = \sum_{k\in\mathrm{pa}(j)} w_{jk}\mathrm{cov}[x_i, x_k] + I_{ij}v_j$$

<u>例 Example</u>



$$\boldsymbol{\mu} = (b_1, b_2 + w_{21}b_1, b_3 + w_{32}b_2 + w_{32}w_{21}b_1)^{\mathrm{T}}$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} v_1 & w_{21}v_1 & w_{32}w_{21}v_1 \\ w_{21}v_1 & v_2 + w_{21}^2 v_1 & w_{32}(v_2 + w_{21}^2 v_1) \\ w_{32}w_{21}v_1 & w_{32}(v_2 + w_{21}^2 v_1) & v_3 + w_{32}^2(v_2 + w_{21}^2 v_1) \end{pmatrix}$$

グラフィカルモデル Graphical models
ガウス変数　Gaussian variables

$$p(x_i|\mathrm{pa}_i) = \mathcal{N}\left(x_i \,\middle|\, \sum_{j\in\mathrm{pa}_i} w_{ij}x_j + b_i, v_i\right)$$

各ノードはガウス変数であり，
その平均は親ノードの線形関数で表される.
Each node is Gaussian and the mean
is a linear function of the parents.

多変量ガウス変数に対しては,
For multivariate Gaussian variables,

$$p(\mathbf{x}_i|\mathrm{pa}_i) = \mathcal{N}\left(\mathbf{x}_i \,\middle|\, \sum_{j\in\mathrm{pa}_i} \mathbf{W}_{ij}\mathbf{x}_j + \mathbf{b}_i, \mathbf{\Sigma}_i\right)$$

グラフィカルモデル Graphical models
二値分類器　Binary classifier

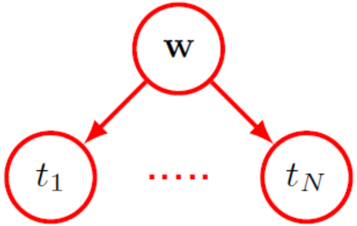
学習パラメータに対するガウス事前分布付き二値分類モデル
Two-class classifier model with Gaussian prior over the learnable parameters

$$p(\mathbf{t}, \mathbf{w}|\mathbf{X}, \lambda) = p(\mathbf{w}|\lambda) \prod_{n=1}^{N} p(t_n|\mathbf{w}, \mathbf{x}_n)$$

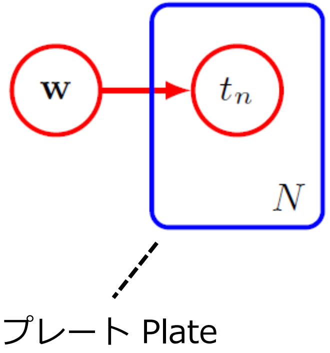
ただし where

$$p(\mathbf{w}|\lambda) = \mathcal{N}(\mathbf{w}|\mathbf{0}, \lambda\mathbf{I})$$

$$p(t|\mathbf{x}, \mathbf{w}) = y(\mathbf{x}, \mathbf{w})^{t} \{1 - y(\mathbf{x}, \mathbf{w})\}^{(1-t)}$$



または or



グラフィカルモデル Graphical models
パラメータと観測　Parameters and observations

学習パラメータに対するガウス事前分布付き二値分類モデル
Two-class classifier model with Gaussian prior over the learnable parameters

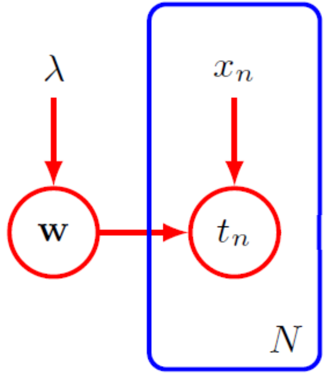
$$p(\mathbf{t}, \mathbf{w}|\mathbf{X}, \lambda) = p(\mathbf{w}|\lambda) \prod_{n=1}^{N} p(t_n|\mathbf{w}, \mathbf{x}_n)$$

ただし where

$$p(\mathbf{w}|\lambda) = \mathcal{N}(\mathbf{w}|\mathbf{0}, \lambda\mathbf{I})$$

$$p(t|\mathbf{x}, \mathbf{w}) = y(\mathbf{x}, \mathbf{w})^t \{1 - y(\mathbf{x}, \mathbf{w})\}^{(1-t)}$$

決定的パラメータが陽に描かれた場合
The deterministic parameters are shown explicitly.



グラフィカルモデル Graphical models
パラメータと観測　Parameters and observations

学習パラメータに対するガウス事前分布付き二値分類モデル
Two-class classifier model with Gaussian prior over the learnable parameters

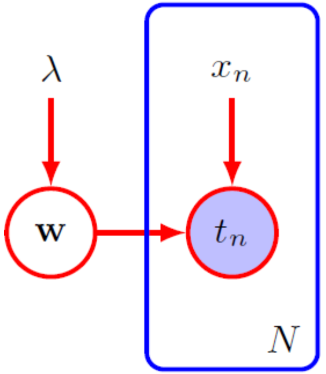
$$p(\mathbf{t}, \mathbf{w}|\mathbf{X}, \lambda) = p(\mathbf{w}|\lambda) \prod_{n=1}^{N} p(t_n|\mathbf{w}, \mathbf{x}_n)$$

ただし where

$$p(\mathbf{w}|\lambda) = \mathcal{N}(\mathbf{w}|\mathbf{0}, \lambda\mathbf{I})$$

$$p(t|\mathbf{x}, \mathbf{w}) = y(\mathbf{x}, \mathbf{w})^t \{1 - y(\mathbf{x}, \mathbf{w})\}^{(1-t)}$$

$\{t_n\}$が観測されている場合
When $\{t_n\}$ are observed,



グラフィカルモデル Graphical models
パラメータと観測　Parameters and observations

学習パラメータに対するガウス事前分布付き二値分類モデル
Two-class classifier model with Gaussian prior over the learnable parameters

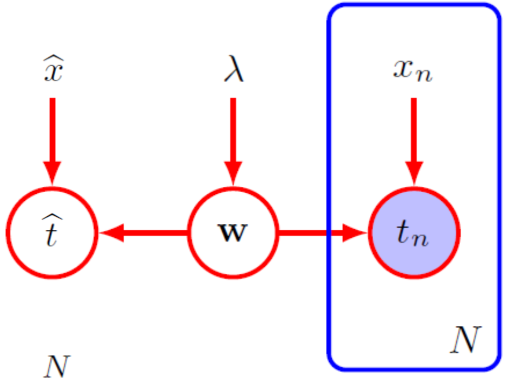
$$p(\mathbf{t}, \mathbf{w}|\mathbf{X}, \lambda) = p(\mathbf{w}|\lambda) \prod_{n=1}^{N} p(t_n|\mathbf{w}, \mathbf{x}_n)$$

ただし where

$$p(\mathbf{w}|\lambda) = \mathcal{N}(\mathbf{w}|\mathbf{0}, \lambda\mathbf{I})$$

$$p(t|\mathbf{x}, \mathbf{w}) = y(\mathbf{x}, \mathbf{w})^t \{1 - y(\mathbf{x}, \mathbf{w})\}^{(1-t)}$$

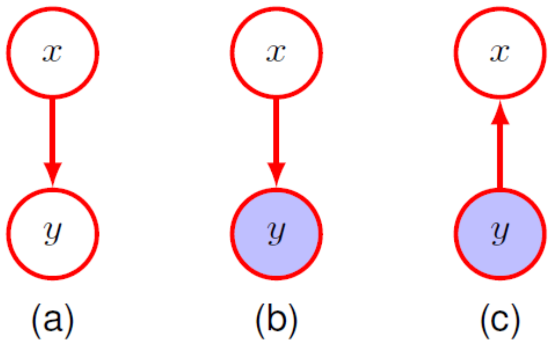
$\{t_n\}$が観測されている場合
When $\{t_n\}$ are observed,



新たな入力 $\widehat{x}$ と対応する $\widehat{t}$ を含む分布
Distribution that contains a new input $\widehat{x}$ and its corresponding $\widehat{t}$

$$p(\widehat{t}, \mathbf{t}, \mathbf{w}|\widehat{x}, \mathbf{X}, \lambda) = p(\mathbf{w}|\lambda)p(\widehat{t}|\mathbf{w}, \widehat{x}) \prod_{n=1}^{N} p(t_n|\mathbf{w}, \mathbf{x}_n)$$

グラフィカルモデル Graphical models
ベイズの定理 Bayes' theorem



(a) $p(x, y) = p(x)p(y|x)$

(b) y を観測したとする.

Suppose we observe y.

(c) $$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

$$p(y) = \sum_{x'} p(y|x')p(x')$$

条件付き独立性 Conditional independence

cが与えられたとき，aがbに対して独立.
a is independent of b given c.

$$p(a|b,c) = p(a|c)$$

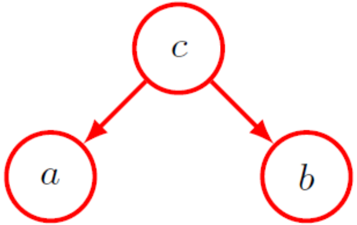
別の表現では Equivalently,

$$\begin{aligned} p(a,b|c) &= p(a|b,c)p(b|c) \\ &= p(a|c)p(b|c) \end{aligned}$$

表記法 Notation $$a \perp\!\!\!\perp b \mid c$$

条件付き独立性 Conditional independence 3つのグラフの例 Three example graphs

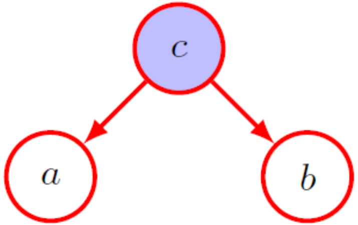
例1 Example 1: tail-to-tail



$$p(a,b,c) = p(a|c)p(b|c)p(c)$$

$$p(a,b) = \sum_c p(a|c)p(b|c)p(c)$$

$$a \not\perp\!\!\!\perp b \mid \emptyset$$

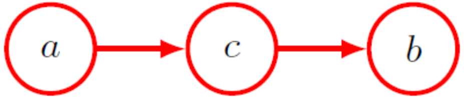


$$\begin{aligned} p(a,b|c) &= \frac{p(a,b,c)}{p(c)} \\ &= p(a|c)p(b|c) \end{aligned}$$

$$a \perp\!\!\!\perp b \mid c$$

条件付き独立性 Conditional independence
3つのグラフの例 Three example graphs

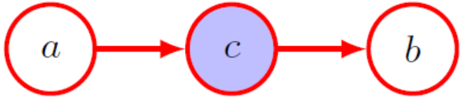
例2 Example 2: head-to-tail



$$p(a,b,c) = p(a)p(c|a)p(b|c)$$

$$p(a,b) = p(a)\sum_c p(c|a)p(b|c) = p(a)p(b|a)$$

$$a \not\perp\!\!\!\perp b \mid \emptyset$$

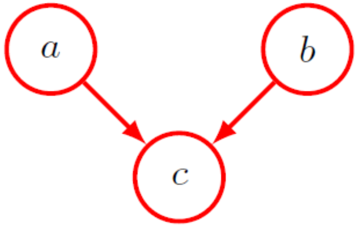


$$\begin{aligned} p(a,b|c) &= \frac{p(a,b,c)}{p(c)} \\ &= \frac{p(a)p(c|a)p(b|c)}{p(c)} \\ &= p(a|c)p(b|c) \end{aligned}$$

$$a \perp\!\!\!\perp b \mid c$$

条件付き独立性 Conditional independence
3つのグラフの例 Three example graphs

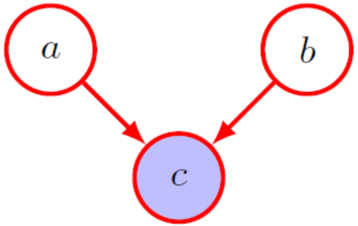
例3 Example 3: head-to-head



$$p(a,b,c) = p(a)p(b)p(c|a,b)$$

$$p(a,b) = p(a)p(b)$$

$$a \perp\!\!\!\perp b \mid \emptyset$$



$$\begin{aligned} p(a,b|c) &= \frac{p(a,b,c)}{p(c)} \\ &= \frac{p(a)p(b)p(c|a,b)}{p(c)} \end{aligned}$$

$$a \not\!\perp\!\!\!\perp b \mid c$$

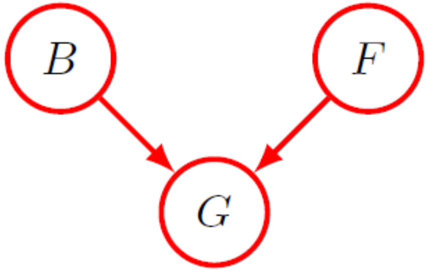
例1，2と反対の結果．This is the opposite of Examples 1 and 2.

条件付き独立性 Conditional independence
「弁明」 Explaining away

燃料切れ？Am I out of fuel?

$$
\begin{aligned}
p(G=1|B=1,F=1) &= 0.8 \\
p(G=1|B=1,F=0) &= 0.2 \\
p(G=1|B=0,F=1) &= 0.2 \\
p(G=1|B=0,F=0) &= 0.1
\end{aligned}
$$

$$
\begin{aligned}
p(B=1) &= 0.9 \qquad p(B=0)=0.1 \\
p(F=1) &= 0.9 \qquad p(F=0)=0.1
\end{aligned}
$$



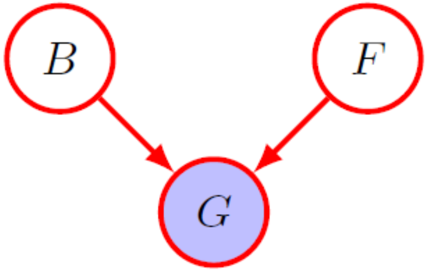
B: バッテリ Battery

F: 燃料タンク Fuel tank

G: 電動燃料系の読み値
Fuel gauge reading

条件付き独立性 Conditional independence 「弁明」 Explaining away

燃料切れ？Am I out of fuel?



$$
\begin{aligned}
p(G=1|B=1,F=1) &= 0.8 \\
p(G=1|B=1,F=0) &= 0.2 \\
p(G=1|B=0,F=1) &= 0.2 \\
p(G=1|B=0,F=0) &= 0.1
\end{aligned}
$$

$$
\begin{aligned}
p(B=1) &= 0.9 \qquad p(B=0)=0.1 \\
p(F=1) &= 0.9 \qquad p(F=0)=0.1
\end{aligned}
$$

B: バッテリ Battery

F: 燃料タンク Fuel tank

G: 電動燃料系の読み値 Fuel gauge reading

$G=0$ を観測することで燃料タンクが空である確率が上昇する.
The probability of an empty fuel tank is increased by observing $G=0$.

$$
\begin{aligned}
p(F=0|G=0) &= \frac{p(G=0|F=0)p(F=0)}{p(G=0)} \\
&\simeq 0.257
\end{aligned}
$$

$$p(G=0|F=0) = \sum_{B\in\{0,1\}} p(G=0|B,F=0)p(B) = 0.81$$

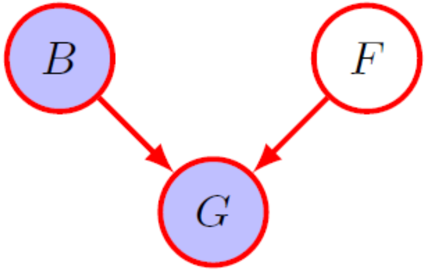
$$p(G=0) = \sum_{B\in\{0,1\}} \sum_{F\in\{0,1\}} p(G=0|B,F)p(B)p(F) = 0.315$$

条件付き独立性 Conditional independence
「弁明」 Explaining away

燃料切れ？Am I out of fuel?

$$
\begin{aligned}
p(G=1|B=1,F=1) &= 0.8\\
p(G=1|B=1,F=0) &= 0.2\\
p(G=1|B=0,F=1) &= 0.2\\
p(G=1|B=0,F=0) &= 0.1
\end{aligned}
$$

$$
\begin{aligned}
p(B=1) &= 0.9 \qquad p(B=0)=0.1\\
p(F=1) &= 0.9 \qquad p(F=0)=0.1
\end{aligned}
$$



B: バッテリ Battery

F: 燃料タンク Fuel tank

G: 電動燃料系の読み値
Fuel gauge reading

$B=0$ を観測することで燃料タンクが空である確率が減少する.
The probability of an empty fuel tank is reduced by observing $B=0$.

$$
\begin{aligned}
p(F=0|G=0,B=0) &= \frac{p(G=0|B=0,F=0)p(F=0)}{\sum_{F\in\{0,1\}} p(G=0|B=0,F)p(F)}\\
&\simeq 0.111
\end{aligned}
$$

条件付き独立性 Conditional independence
有向分離（D分離） D-separation

有向グラフ内のA, B およびC を互いに重複しないノード集合とする.
A, B, and C are non-intersecting subsets of nodes in a directed graph.

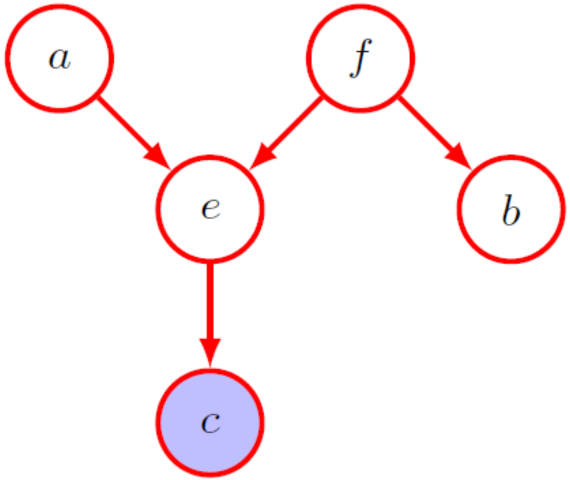
A からB への経路は，以下のいずれかを満たすノードを含むとき，遮断されているという.
A path from A to B is blocked if it contains a node such that either

- 集合C に含まれるノードであって，経路に含まれる矢印がそこでhead-to-tailあるいはtail-to-tailである.
  the arrows on the path meet either head-to-tail or tail-to-tail at the node, and the node is in the set C,
  または or
- 経路に含まれる矢印がそのノードでhead-to-headであり，自身あるいはそのすべての子孫のいずれもが集合Cに含まれない.
  the arrows meet head-to-head at the node, and neither the node, nor any of its descendants, are in the set C.

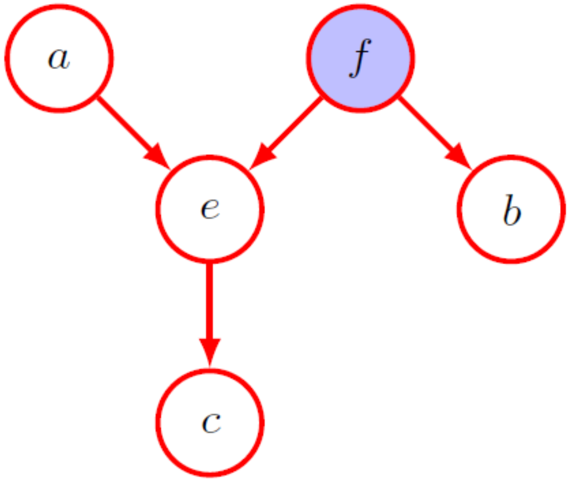
A からB へのすべての経路が遮断されているとき，A はC によってB から有向分離されているという.
If all paths from A to B are blocked, A is said to be d-separated from B by C.

A がC によってB から有向分離されていれば，グラフの全変数上の同時分布は$A \perp\!\!\!\perp B \mid C$ を満たす.
If A is d-separated from B by C, the joint distribution over all variables in the graph satisfies $A \perp\!\!\!\perp B \mid C$.

条件付き独立性 Conditional independence
有向分離（D分離） D-separation

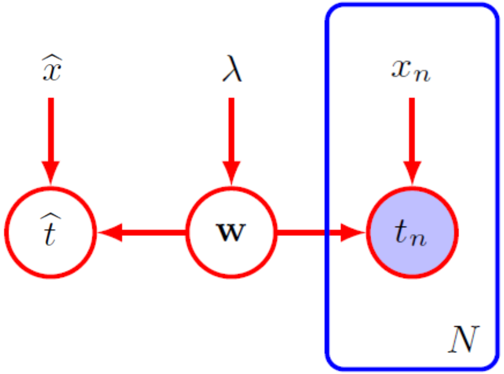


$$a \not\perp\!\!\!\perp b \mid c$$



$$a \perp\!\!\!\perp b \mid f$$

条件付き独立性 Conditional independence
有向分離（D分離） D-separation



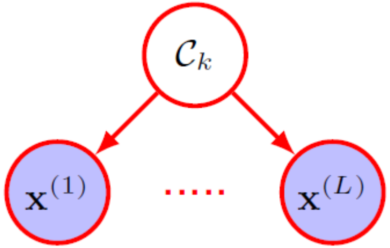
$\{t_1, \ldots, t_N\}$: i.i.d.データ i.i.d. data

ノード $\widehat{t}$ から t_n のうちの任意のノードへの経路に関してノード $\mathbf{w}$ はtail-to-tailである.
The node for $\mathbf{w}$ is tail-to-tail with respect to the path from $\widehat{t}$ to any one of the nodes t_n.

⇨ $\widehat{t} \perp\!\!\!\perp t_n \mid \mathbf{w}$

条件付き独立性 Conditional independence
ナイーブベイズ Naive Bayes

ナイーブベイズモデル Naive Bayes model



$$p(\mathbf{x}|\mathcal{C}_k) = \prod_{l=1}^{L} p(\mathbf{x}^{(l)}|\mathcal{C}_k)$$

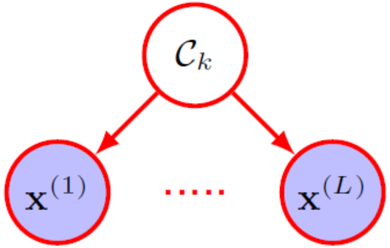
$$\mathbf{x} = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(L)})$$

$\mathcal{C}_k$ を観測すると, $\mathbf{x}^{(i)}$ と $\mathbf{x}^{(j)}$ ($j \neq i$) との間の経路が遮断される.

An observation of $\mathcal{C}_k$ would block the path between $\mathbf{x}^{(i)}$ and $\mathbf{x}^{(j)}$ for $j \neq i$.

条件付き独立性 Conditional independence
ナイーブベイズ Naive Bayes

ナイーブベイズモデル Naive Bayes model



$$p(\mathbf{x}|\mathcal{C}_k) = \prod_{l=1}^{L} p(\mathbf{x}^{(l)}|\mathcal{C}_k)$$

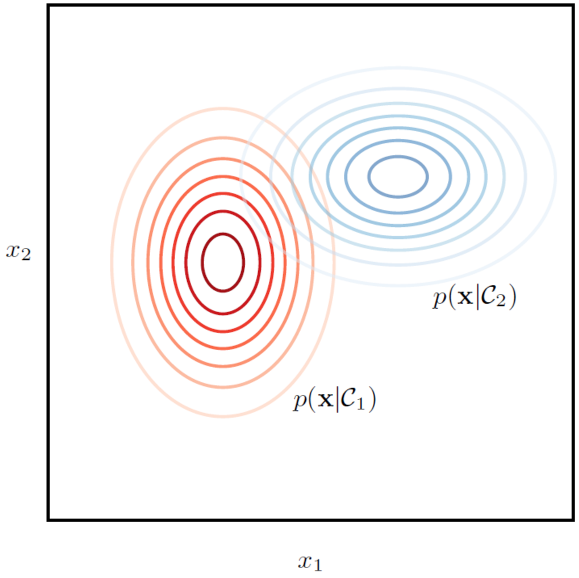
$$\mathbf{x} = (\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(L)})$$

ベクトル $\mathbf{x}$ がクラス $\mathcal{C}_k$ に所属する確率
The probability that a vector $\mathbf{x}$ belongs to class $\mathcal{C}_k$

$$p(\mathcal{C}_k|\mathbf{x}) = \frac{p(\mathbf{x}|\mathcal{C}_k)p(\mathcal{C}_k)}{p(\mathbf{x})}$$ ただし where $$p(\mathbf{x}) = \sum_{k=1}^{K} p(\mathbf{x}|\mathcal{C}_k)p(\mathcal{C}_k)$$

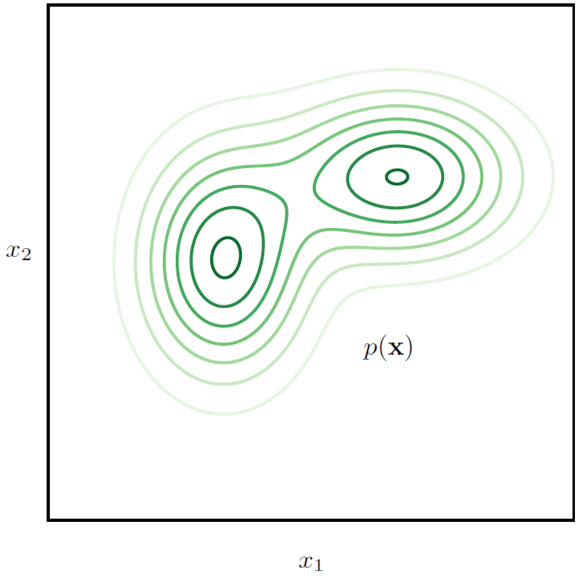
条件付き独立性 Conditional independence
ナイーブベイズ Naive Bayes

Factorized 分解されている



$$p(\mathbf{x}|\mathcal{C}_k) = p(x_1|\mathcal{C}_k)p(x_2|\mathcal{C}_k)$$

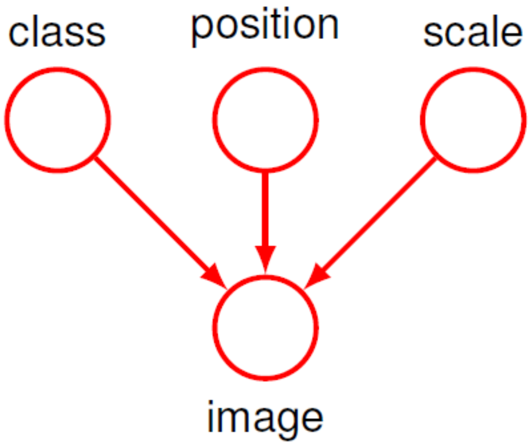
Not factorized 分解されていない



$$p(\mathbf{x}) = \sum_{k=1}^{K} p(\mathbf{x}|\mathcal{C}_k)p(\mathcal{C}_k)$$

$$\mathbf{x} = (x_1, x_2)$$

条件付き独立性 Conditional independence
生成モデル Generative models

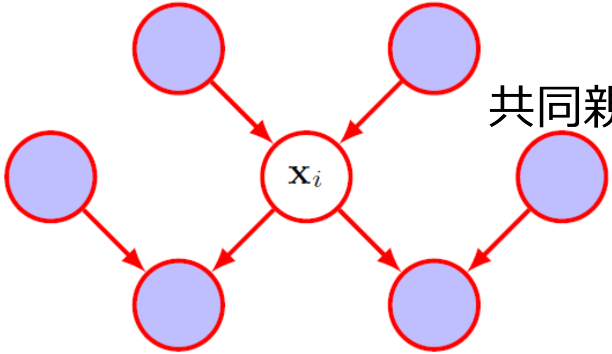


物体の像が生成される過程を表現するグラフィカルモデル
A graphical model representing the process by which images of objects are created

条件付き独立性 Conditional independence
マルコフブランケット Markov blanket

マルコフブランケットに含まれるノード
Nodes included in the Markov blanket

親 Parents



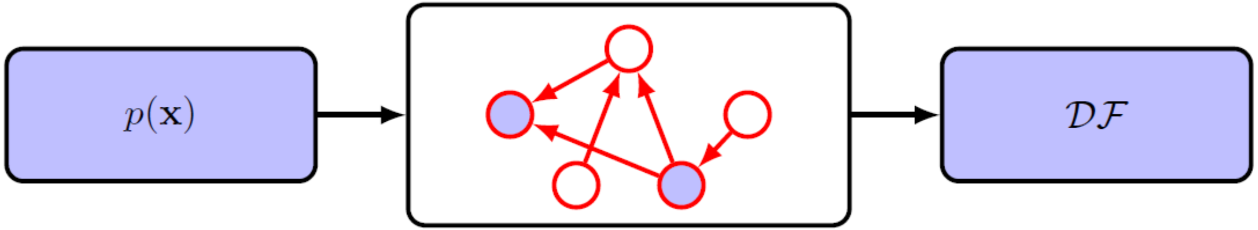
共同親 Co-parents

子 Children

$$p(\mathbf{x}_i|\mathbf{x}_{\{j\neq i\}}) = \frac{p(\mathbf{x}_1,\ldots,\mathbf{x}_M)}{\int p(\mathbf{x}_1,\ldots,\mathbf{x}_M)\,\mathrm{d}\mathbf{x}_i} = \frac{\prod_k p(\mathbf{x}_k|\mathrm{pa}_k)}{\int \prod_k p(\mathbf{x}_k|\mathrm{pa}_k)\,\mathrm{d}\mathbf{x}_i}$$

$\mathbf{x}_i$ に依存しない因子は分子と分母とでキャンセルされる.
Factors independent of $\mathbf{x}_i$ cancel between numerator and denominator.

条件付き独立性 Conditional independence
フィルタとしてのグラフ Graph as filters



分解特性 Factorization property

$$p(x_1, \ldots, x_K) = \prod_{k=1}^{K} p(x_k | \mathrm{pa}(k))$$

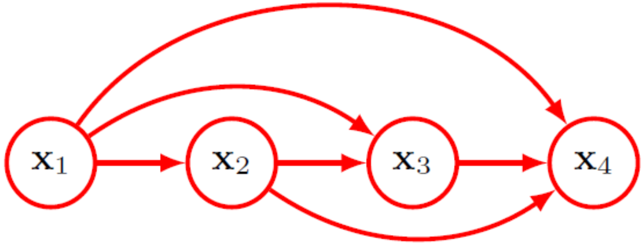
系列モデル Sequence models

<u>系列データ Sequential data</u>

例：マイクからのオーディオ信号，ある地点における日々の雨量観測値
Examples: audio signals from a microphone,
daily rain fall measurements at a particular location

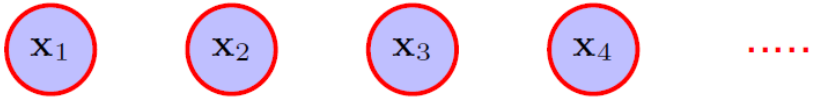
以下は，一般性を完全に保っているときの表現
The following expression has complete generality.

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_N) = p(\mathbf{x}_1) \prod_{n=2}^{N} p(\mathbf{x}_n | \mathbf{x}_1, \ldots, \mathbf{x}_{n-1})$$



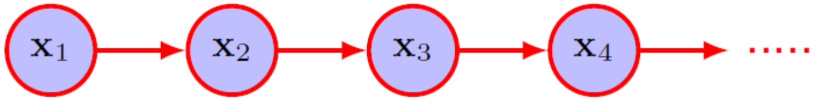
系列モデル Sequence models

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \prod_{n=1}^{N} p(\mathbf{x}_n)$$



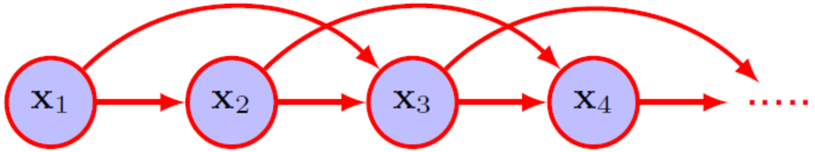
順序情報を無視
Ignore the ordering information

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_N) = p(\mathbf{x}_1) \prod_{n=2}^{N} p(\mathbf{x}_n | \mathbf{x}_{n-1})$$



一次マルコフ連鎖
First-order Markov chain

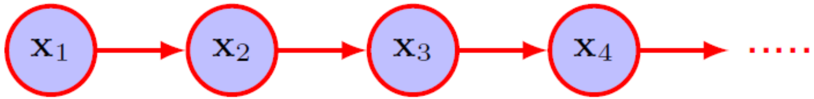
$$p(\mathbf{x}_1, \ldots, \mathbf{x}_N) = p(\mathbf{x}_1) p(\mathbf{x}_2 | \mathbf{x}_1) \prod_{n=3}^{N} p(\mathbf{x}_n | \mathbf{x}_{n-1}, \mathbf{x}_{n-2})$$



二次マルコフ連鎖
Second-order Markov chain

系列モデル Sequence models

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_N) = p(\mathbf{x}_1) \prod_{n=2}^{N} p(\mathbf{x}_n | \mathbf{x}_{n-1})$$



一次マルコフ連鎖
First-order Markov chain

有向分離を用いると，Using d-separation,

$$p(\mathbf{x}_n | \mathbf{x}_1, \ldots, \mathbf{x}_{n-1}) = p(\mathbf{x}_n | \mathbf{x}_{n-1})$$

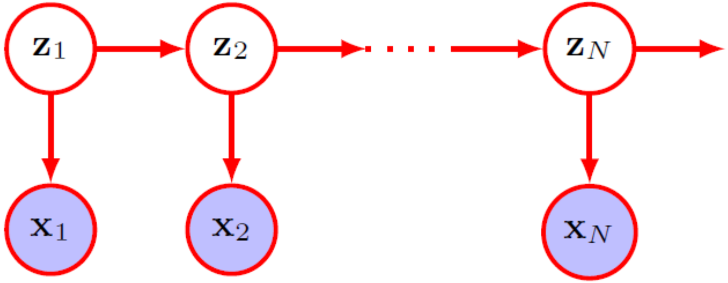
観測変数が K 個の状態をもつ離散変数であるとき，
If the observed variables are discrete variables having K states,

パラメータ数 The number of parameters: $K(K-1)$

M 次マルコフ連鎖では
In an M-th order Markov chain, $K^M(K-1)$

系列モデル Sequence models
隠れ変数 Hidden variables

状態空間モデル State-space model



潜在変数 Latent variables

観測変数 Observed variables

⇨ $\mathbf{z}_{n+1} \perp\!\!\!\perp \mathbf{z}_{n-1} \mid \mathbf{z}_n$

$\mathbf{x}_{n+1}$ は過去のすべての観測変数に依存する.
$\mathbf{x}_{n+1}$ depends on all previous observed variables.

上のモデルの同時分布 The joint distribution of the above model

$$p(\mathbf{x}_1, \ldots, \mathbf{x}_N, \mathbf{z}_1, \ldots, \mathbf{z}_N) = p(\mathbf{z}_1) \left[\prod_{n=2}^{N} p(\mathbf{z}_n | \mathbf{z}_{n-1}) \right] \prod_{n=1}^{N} p(\mathbf{x}_n | \mathbf{z}_n)$$

提出課題 IV：グラフィカルモデル
Assignment IV: Graphical Models

提出期限：**11月12日（火曜日）23:59:00** **[日本標準時]**
Submission deadline: **November 12 (Tuesday) 23:59:00** **[Japan Standard Time]**

提出課題は「一般」チャネルの「ファイル」にアップロードされます．
同チャネルに出現する通知のリンク先から解答を送信（提出）してください．
Assignments will be uploaded to "File" in the "General" channel. Send (submit) your answers via the link that will appear in the same channel.

- 全6回の課題への解答をもとに成績評価を行います．
  Your grades will be based on your answers to all six assignments.
- 解答送信後の解答再送信はできません．
  Once submitted, answers cannot be resubmitted.
- 提出期限は一切延長しません．
  The submission deadline will never be extended.