情報通信システム(特)論II

Information and Communication Systems II

第7回講義資料

Lecture notes 7

確率伝播法

Belief Propagation

豊橋技術科学大学

Toyohashi University of Technology

電気・電子情報工学系

Department of Electrical and Electronic Information Engineering

准教授　竹内啓悟

Associate Professor　Keigo Takeuchi

無線通信路(Wireless channels)

物理層における無線通信システムのほとんどは、物理法則に基づいて以下のシステムでモデル化できる。

Almost all wireless communication systems in the physical layer can be modeled with the following system based on physical laws.

無線通信路の離散時間統計モデル(Discrete-time statistical model of wireless channels)

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{w}, \qquad \boldsymbol{w} \sim \mathcal{CN}(\boldsymbol{0}, N_0 \boldsymbol{I}_M)$$

$\boldsymbol{x} \in \mathbb{C}^N$: 送信ベクトル(Transmitted vector)　$\boldsymbol{y} \in \mathbb{C}^M$: 受信ベクトル(Received vector)

$\boldsymbol{H} \in \mathbb{C}^{M \times N}$: 通信路行列(Channel matrix)

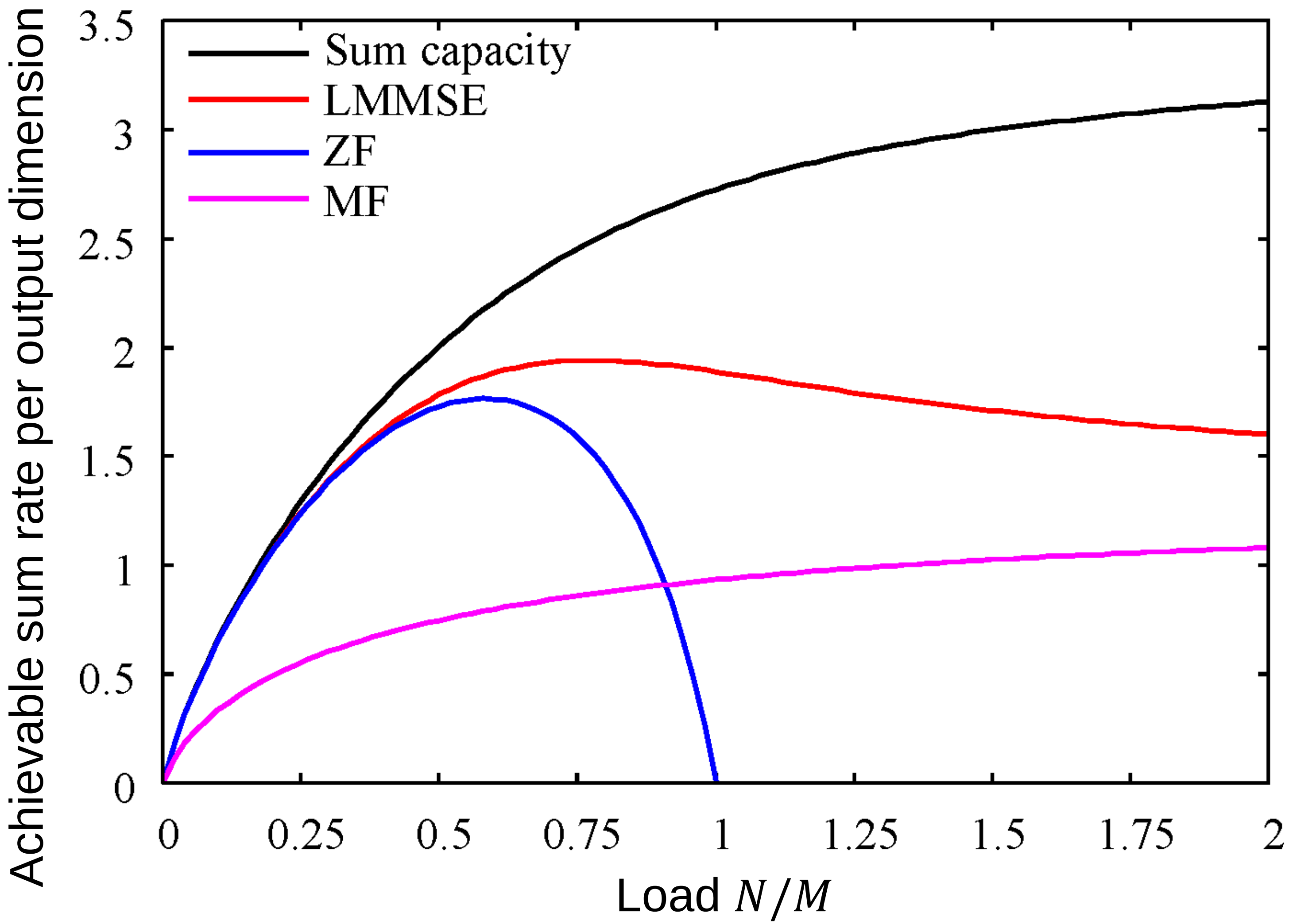
$\boldsymbol{w} \sim \mathcal{CN}(\boldsymbol{0}, N_0 \boldsymbol{I}_M)$: 加法的白色ガウス雑音(AWGN)ベクトル

Additive white Gaussian noise (AWGN) vector

統計的仮定(Statistical assumptions)

- 送信ベクトル$\boldsymbol{x}$は平均$\boldsymbol{0}$である。(The transmitted vector $\boldsymbol{x}$ is zero mean.)
- 確率変数$\boldsymbol{H}$、$\boldsymbol{x}$、$\boldsymbol{w}$は独立である。($\boldsymbol{H}$, $\boldsymbol{x}$, and $\boldsymbol{w}$ are independent random variables.)

達成可能な総和レートの比較(Comparisons for achievable sum rates)



i.i.d. Rayleigh fading, $N \to \infty$, and $P/N_0 = 10$ dB.

事後平均推定量(Posterior mean estimator)

各データシンボルx_nは有限個の信号点集合$\mathcal{M} \subset \mathbb{C}$上で値を取るとする。

Suppose that each data symbol x_n takes values on a finite constellation $\mathcal{M} \subset \mathbb{C}$.

事後平均推定量(Posterior mean estimator)

$$\mathbb{E}[\boldsymbol{x}|\boldsymbol{H},\boldsymbol{y}] = \sum_{\boldsymbol{x}\in\mathcal{M}^N} \boldsymbol{x}\frac{p(\boldsymbol{y}|\boldsymbol{H},\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{y}|\boldsymbol{H})} = \frac{\sum_{\boldsymbol{x}\in\mathcal{M}^N}\boldsymbol{x}p(\boldsymbol{y}|\boldsymbol{H},\boldsymbol{x})p(\boldsymbol{x})}{\sum_{\boldsymbol{x}\in\mathcal{M}^N}p(\boldsymbol{y}|\boldsymbol{H},\boldsymbol{x})p(\boldsymbol{x})}.$$

計算量(Complexity)

直接計算すると、少なくとも$O(|\mathcal{M}|^N)$回の足し算が必要である。

Direct computation requires at least $O(|\mathcal{M}|^N)$ addtions.

計算量の少ない近似計算法はあるか？

Is there an approximate and low-complexity method?

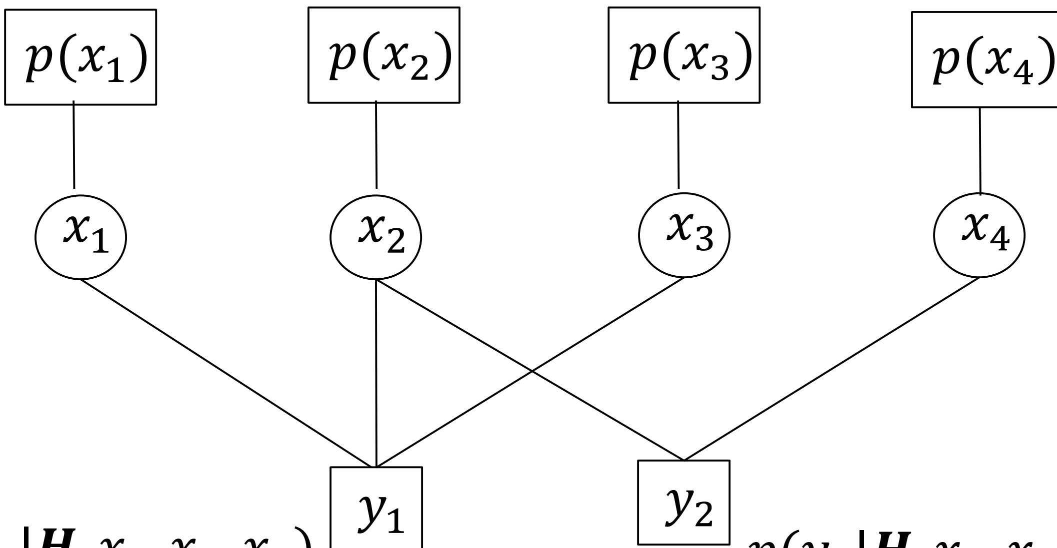
ファクターグラフ表現(Factor graph representation)

データベクトル$\boldsymbol{x}$の要素は独立と仮定する。

Assume the independence of the elements of the data vector $\boldsymbol{x}$.

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & 0 \\ 0 & h_{22} & 0 & h_{24} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

ファクターグラフ(Factor graph)



$p(y_1|\boldsymbol{H}, x_1, x_2, x_3)$ $p(y_2|\boldsymbol{H}, x_2, x_4)$

ファクターグラフの書き方(How to draw a factor graph)

1. 推定すべき未知の確率変数を変数ノードと呼び、丸で書く。事前分布や既知の確率変数によって決まる制約を表すノードをファクターノードと呼び、四角で書く。

   Unknown random variables to be estimated are called variable nodes, denoted by circles. Nodes representing constraints due to prior distributions or known random variables are referred to as factor nodes, denoted by squares.

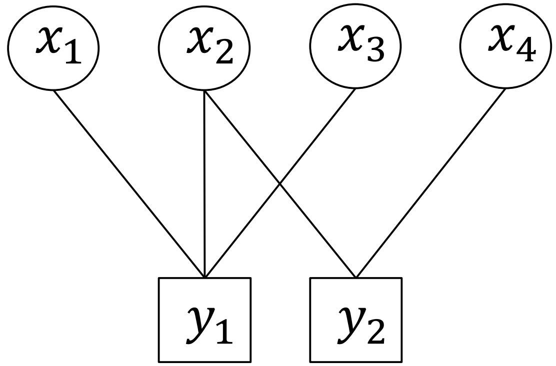
2. 通信路行列の要素$h_{m,n}$が0でないとき、ファクターノードy_mと変数ノードx_nを接続する。

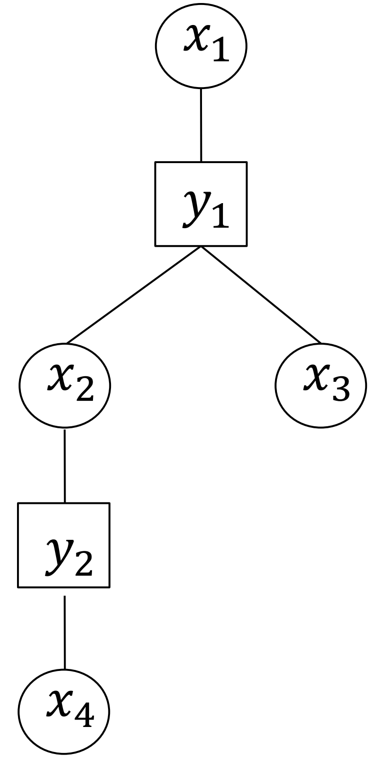
   When an element $h_{m,n}$ of the channel matrix is non-zero, the factor node y_m is connected to the variable node x_n.

注意(Remark)

事前分布を表すファクターノードは省略される場合がある。
Factor nodes representing prior distributions may be omitted.

木構造のファクターグラフ(Factor graph with tree structure)





サイクルがない。

There are no cycles.

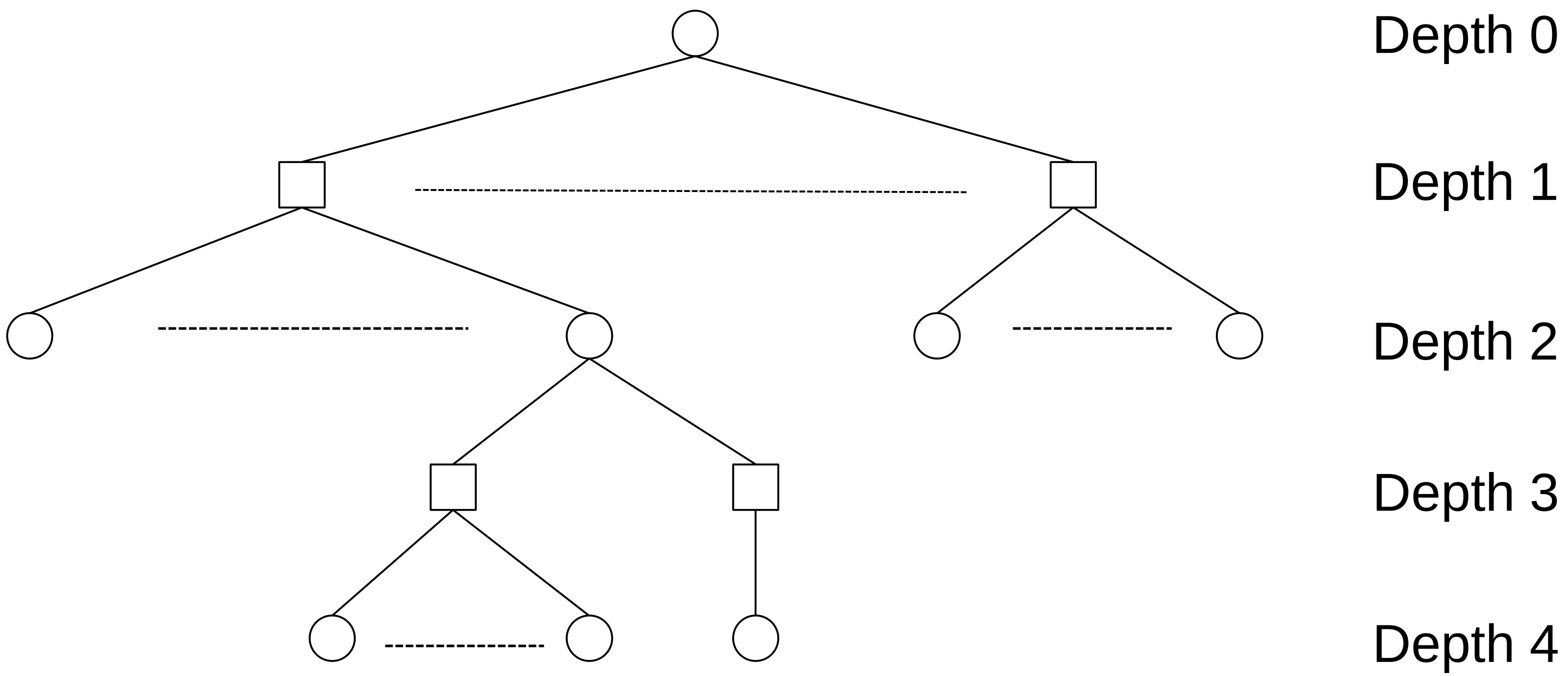
確率構造(Probabilistic structure)

$$p(\boldsymbol{y}|\boldsymbol{H},\boldsymbol{x}) = p(y_1|\boldsymbol{H},x_1,x_2,x_3)p(y_2|\boldsymbol{H},x_2,x_4).$$

y_2が$p(y_1|\boldsymbol{H},x_1,x_2,x_3)$で出てきた$x_1$や$x_3$に依存しないことが重要である。

It is important that y_2 is independent of x_1 and x_3, which appeared in $p(y_1|\boldsymbol{H},x_1,x_2,x_3)$.

木構造のファクターグラフ(Factor graph with tree structure)



性質(Properties)

- ファクターノードは子ノードを持つ。(Any factor node has child nodes.)
- 根ノードを除いて、各ノードは親ノードを一つ持つ。

  With the exception of the root node, any node has one parent node

局所構造(Local structure)

ファクターグラフは深さ$2D$の木であるとしよう。

Suppose that the factor graph is a tree with depth $2D$.

$\mathcal{X}_i = \{x_{i,1}, \dots, x_{i,n_i}\}$：深さ$2i$にある変数ノード(Variable nodes in depth $2i$.)

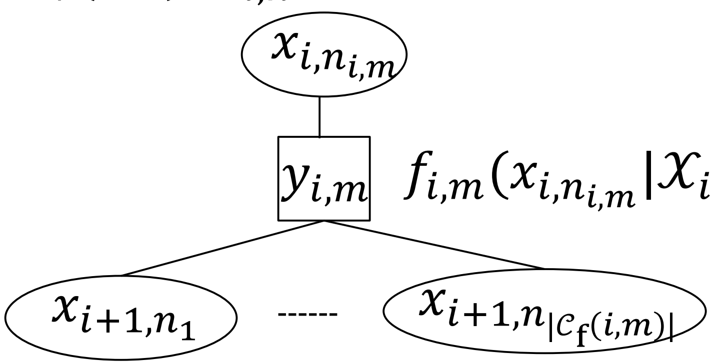
$\mathcal{Y}_i = \{y_{i,1}, \dots, y_{i,m_i}\}$：深さ$2i+1$にあるファクターノード(Factor nodes in depth $2i+1$.)

$n_{i,m}$：$y_{i,m}$の親を指定する番号(Index indicating the parent of $y_{i,m}$)

$m_{i,n}$：$x_{i,n}$の親を指定する番号(Index indicating the parent of $x_{i,n}$)

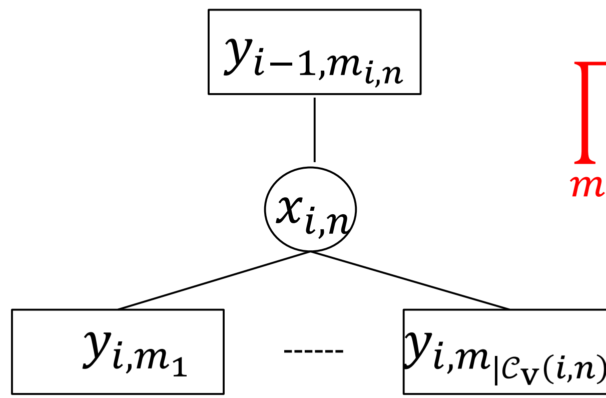
$\mathcal{C}_{\mathrm{f}}(i,m)$：$y_{i,m}$の子供を指定する番号集合(Set of indices indicating the children of $y_{i,m}$)

$\mathcal{C}_{\mathrm{v}}(i,n)$：$x_{i,n}$の子供を指定する番号集合(Set of indices indicating the children of $x_{i,n}$)



$f_{i,m}(x_{i,n_{i,m}}|\mathcal{X}_{i+1,m})$

$\mathcal{X}_{i+1,m} = \{x_{i+1,n} : n \in \mathcal{C}_{\mathrm{f}}(i,m)\}$



$$\prod_{m\in\emptyset} f_{i,m} = 1$$

$$p(\boldsymbol{y}|\boldsymbol{H},\boldsymbol{x}) \stackrel{\text{def}}{=} \prod_{i=0}^{D-1}\prod_{m=1}^{m_i} f_{i,m}\left(x_{i,n_{i,m}}|\mathcal{X}_{i+1,m}\right) = \prod_{i=0}^{D-1}\prod_{n=1}^{n_i}\prod_{m\in\mathcal{C}_{\mathrm{v}}(i,n)} f_{i,m}\left(x_{i,n}|\mathcal{X}_{i+1,m}\right).$$

周辺事後分布の計算(Computation of marginal posterior distributions)

一般性を失うことなく、根ノード$x_{0,1}$の周辺事後分布を計算する。

Without loss of generality, we compute the marginal posterior distribution of the root node $x_{0,1}$.

$$p(x_{0,1}|\boldsymbol{H},\boldsymbol{y}) = \sum_{\backslash x_{0,1}} p(\boldsymbol{x}|\boldsymbol{H},\boldsymbol{y}) \propto \sum_{\backslash x_{0,1}} p(\boldsymbol{y}|\boldsymbol{H},\boldsymbol{x}) \prod_{n=1}^{N} p(x_n).$$

総和は$x_{0,1}$を固定した全組み合わせ$\{\boldsymbol{x} \in \mathcal{M}^N : x_{0,1} \text{ fixed}\}$に関して取られる。

The summation is over all possible combinations $\{\boldsymbol{x} \in \mathcal{M}^N : x_{0,1} \text{ fixed}\}$ for fixed $x_{0,1}$.

木であるための条件(Tree condition)

$i \neq 0$に対して$\{\mathcal{X}_{i,m} : \text{for all } m\}$は$\mathcal{X}_i$の分割である。

$\{\mathcal{X}_{i,m} : \text{for all } m\}$ is a disjoint partition of $\mathcal{X}_i$ for $i \neq 0$.

$$\cup_m \mathcal{X}_{i,m} = \mathcal{X}_i, \qquad \mathcal{X}_{i,m} \cap \mathcal{X}_{i,m'} = \emptyset \ \text{ for } m \neq m'.$$

$p(\boldsymbol{y}|\boldsymbol{H},\boldsymbol{x})$の因子分解を使って、(Using the factor decomposition of $p(\boldsymbol{y}|\boldsymbol{H},\boldsymbol{x})$ yields)

$$p(x_{0,1}|\boldsymbol{H},\boldsymbol{y}) \propto p(x_{0,1}) \sum_{\backslash x_{0,1}} \prod_{i=0}^{D-1} \prod_{n=1}^{n_i} \prod_{m \in \mathcal{C}_\mathrm{v}(i,n)} \{f_{i,m}(x_{i,n}|\mathcal{X}_{i+1,m})\, p(\mathcal{X}_{i+1,m})\}$$

確率伝播法(Belief propagation (BP))

ファクターノード$y_{i,m}$から変数ノード$x_{i,n_{i,m}}$へのメッセージ$q^{D-i}_{m\to n_{i,m}}(x_{i,n_{i,m}})$

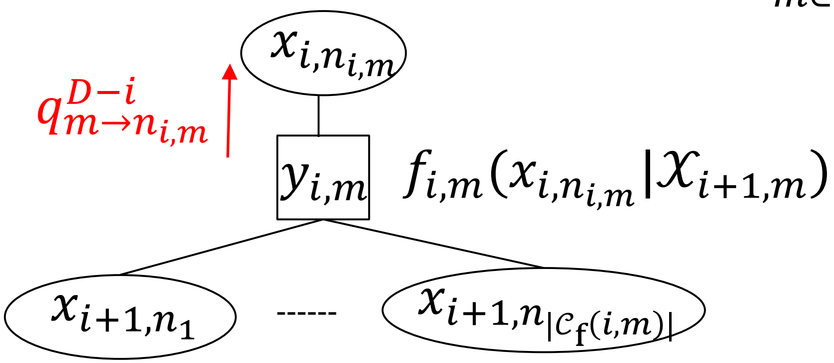
Message $q^{D-i}_{m\to n_{i,m}}(x_{i,n_{i,m}})$ passed from the factor node $y_{i,m}$ to the variable node $x_{i,n_{i,m}}$

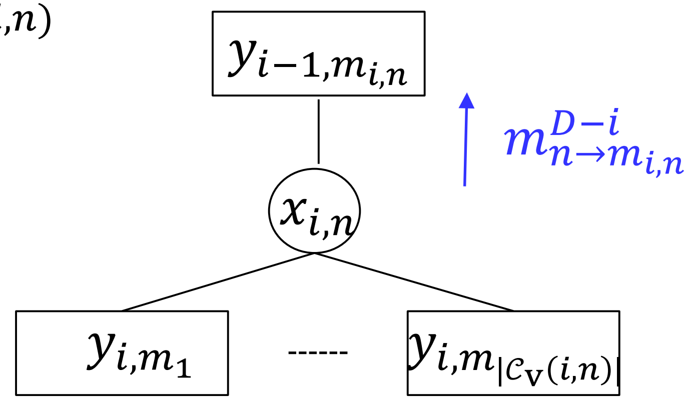
$$q^{D-i}_{m\to n_{i,m}}(x_{i,n_{i,m}}) = \sum_{\mathcal{X}_{i+1,m}} f_{i,m}(x_{i,n_{i,m}}|\mathcal{X}_{i+1,m}) \prod_{n\in\mathcal{C}_{\mathrm{f}}(i,m)} m^{D-(i+1)}_{n\to m}(x_{i+1,n}).$$

変数ノード$x_{i,n}$からファクターノード$y_{i-1,m_{i,n}}$へのメッセージ$m^{D-i}_{n\to m_{i,n}}(x_{i,n})$

Message $m^{D-i}_{n\to m_{i,n}}(x_{i,n})$ passed from the variable node $x_{i,n}$ to the variable node $y_{i-1,m_{i,n}}$

$$m^{D-i}_{n\to m_{i,n}}(x_{i,n}) \propto p(x_{i,n}) \prod_{m\in\mathcal{C}_{\mathrm{v}}(i,n)} q^{D-i}_{m\to n}(x_{i,n}).$$





メッセージは葉から根に向かって計算される。

The messages are computed from the leafs toward the root.

周辺事後分布の計算(Computation of marginal posterior distributions)

定理7.1(Theorem 7.1)

$$p\left(x_{0,1}\middle|\boldsymbol{H},\boldsymbol{y}\right) \propto p\left(x_{0,1}\right) \prod_{m\in\mathcal{C}_{\mathrm{v}}(0,1)} q_{m\to 1}^{D}\left(x_{0,1}\right).$$

証明(Proof)　深さDに関する帰納法によって証明する。

The proof is by induction with respect to the depth D.

$D=1$の場合

Case of $D = 1$

$$p(x_n|\boldsymbol{H},\boldsymbol{y}) \propto p(x_{0,1}) \sum_{\backslash x_{0,1}} \prod_{m\in\mathcal{C}_{\mathrm{v}}(0,1)} \{f_{0,m}\left(x_{0,1}\middle|\mathcal{X}_{1,m}\right) p(\mathcal{X}_{1,m})\}$$

$$= p\left(x_{0,1}\right) \prod_{m\in\mathcal{C}_{\mathrm{v}}(0,1)} \sum_{\mathcal{X}_{1,m}} f_{0,m}\left(x_{0,1}\middle|\mathcal{X}_{1,m}\right) \prod_{n\in\mathcal{C}_{\mathrm{f}}(0,m)} p\left(x_{1,n}\right)$$

$$= p\left(x_{0,1}\right) \prod_{m\in\mathcal{C}_{v}(0,1)} q_{m\to 1}^{1}\left(x_{0,1}\right).$$

$$\because \quad q_{m\to 1}^{1}\left(x_{0,1}\right) = \sum_{\mathcal{X}_{1,m}} f_{0,m}\left(x_{0,1}\middle|\mathcal{X}_{1,m}\right) \prod_{n\in\mathcal{C}_{\mathrm{f}}(0,m)} m_{n\to m}^{0}\left(x_{1,n}\right),$$

$$p\left(x_{1,n}\right) \propto m_{n\to m}^{0}\left(x_{1,n}\right).$$

定理7.1の証明(Proof of Theorem 7.1)

$D=d$の場合に定理は正しいと仮定して、$D=d+1$の場合を示す。

We prove the case $D=d+1$ under the assumption of the correctness for $D=d$.

$D=1$の場合と同様に、$\mathcal{X}_{d+1}$に関する総和を計算する。

We repeat the proof for $D=1$ to compute the summation over $\mathcal{X}_{d+1}$

$$p(x_{0,1}|\boldsymbol{H},\boldsymbol{y}) \propto p(x_{0,1}) \sum_{\backslash x_{0,1}} \prod_{i=0}^{d} \prod_{n=1}^{n_i} \prod_{m\in\mathcal{C}_{\mathrm{v}}(i,n)} \{f_{i,m}(x_{i,n}|\mathcal{X}_{i+1,m})\, p(\mathcal{X}_{i+1,m})\}$$

$$\propto p(x_{0,1}) \sum_{\backslash x_{0,1}} \prod_{i=0}^{\color{red}{d-1}} \prod_{n=1}^{n_i} \prod_{m\in\mathcal{C}_{\mathrm{v}}(i,n)} \{f_{i,m}(x_{i,n}|\mathcal{X}_{i+1,m})\, p(\mathcal{X}_{i+1,m})\}$$

$$\cdot \prod_{n'=1}^{n_d} \prod_{m'\in\mathcal{C}_{\mathrm{v}}(d,n')} \color{blue}{q^1_{m'\to n'}(x_{d,n'})}.$$

$$\because \quad q^1_{m\to n}(x_{d,n}) = \sum_{\mathcal{X}_{d+1,m}} f_{d,m}(x_{d,n}|\mathcal{X}_{d+1,m}) \prod_{n'\in\mathcal{C}_{\mathrm{f}}(d,m)} m^0_{n'\to m}(x_{d+1,n'}),$$

$$m^0_{n'\to m}(x_{d+1,n'}) \propto p(x_{d+1,n'}).$$

定理7.1の証明(Proof of Theorem 7.1)

$D=d$の場合に帰着する。(We reduce the formula to the case of $D=d$.)

$\{\mathcal{X}_{i,m}\}$は$\mathcal{X}_i$の分割なので、(Since $\{\mathcal{X}_{i,m}\}$ is a disjoint partition of $\mathcal{X}_i$, we have)

$$\prod_{n=1}^{n_{d-1}} \prod_{m\in\mathcal{C}_{\mathrm{v}}(d-1,n)} p(\mathcal{X}_{d,m}) = \prod_{n=1}^{n_{d-1}} \prod_{m\in\mathcal{C}_{\mathrm{v}}(d-1,n)} \prod_{n\in\mathcal{C}_f(d-1,m)} p(x_{d,n}) = \prod_{n=1}^{n_d} p(x_{d,n}).$$

それゆえ、(Thus,)

$$p(x_{0,1}|\boldsymbol{H},\boldsymbol{y}) \propto p(x_{0,1}) \sum_{\backslash x_{0,1}} \prod_{i=0}^{d-2} \prod_{n=1}^{n_i} \prod_{m\in\mathcal{C}_{\mathrm{v}}(i,n)} \{f_{i,m}(x_{i,n}|\mathcal{X}_{i+1,m})\, p(\mathcal{X}_{i+1,m})\}$$

$$\cdot \prod_{n=1}^{n_{d-1}} \prod_{m\in\mathcal{C}_{\mathrm{v}}(d-1,n)} f_{d-1,m}(x_{d-1,n}|\mathcal{X}_{d,m}) \cdot \prod_{n'=1}^{n_d} m^1_{n'\to m_{d,n'}}(x_{d,n'}).$$

$$\because \quad m^1_{n'\to m_{d,n'}}(x_{d,n'}) \propto p(x_{d,n'}) \prod_{m'\in\mathcal{C}_{\mathrm{v}}(d,n')} q^1_{m'\to n'}(x_{d,n'}).$$

定理7.1の証明(Proof of Theorem 7.1)

$\tilde{p}(x_{d,n}) = m^1_{n\to m_{d,n}}(x_{d,n})$を深さ$2d$にある変数ノードの事前分布とみなすと、

Regarding $\tilde{p}(x_{d,n}) = m^1_{n\to m_{d,n}}(x_{d,n})$ as the prior distributions of the variable nodes in depth $2d$ yileds

$$p(x_{0,1}|\boldsymbol{H},\boldsymbol{y}) \propto p(x_{0,1}) \sum_{\backslash x_{0,1}} \prod_{i=0}^{d-1} \prod_{n=1}^{n_i} \prod_{m\in\mathcal{C}_{\mathrm{v}}(i,n)} \{f_{i,m}(x_{i,n}|\mathcal{X}_{i+1,m})\, p(\mathcal{X}_{i+1,m})\},$$

ただし、$p(\mathcal{X}_{d,m}) = \prod_{n\in\mathcal{C}_{\mathrm{f}}(d,m)} \tilde{p}(x_{d,n})$と再定義する。

where $p(\mathcal{X}_{d,m}) = \prod_{n\in\mathcal{C}_{\mathrm{f}}(d,m)} \tilde{p}(x_{d,n})$ is re-defined.

$$\because \quad \prod_{n'=1}^{n_d} m^1_{n'\to m_{d,n'}}(x_{d,n'}) = \prod_{n=1}^{n_{d-1}} \prod_{m\in\mathcal{C}_{\mathrm{v}}(d-1,n)} \tilde{p}(\mathcal{X}_{d,m}).$$

それゆえ、定理は帰納法の仮定から従う。 ■

Thus, the theorem follows from the induction hypothesis.

確率伝播法の計算量(Complexity of BP)

最も支配的な計算はファクターノードから変数ノードに送るメッセージの更新である。

The most dominant computation is the update of messages passed from each factor node to variable nodes.

$$q_{m\to n_{i,m}}^{D-i}\left(x_{i,n_{i,m}}\right)=\sum_{\mathcal{X}_{i+1,m}} f_{i,m}\left(x_{i,n_{i,m}}\middle|\mathcal{X}_{i+1,m}\right)\prod_{n\in\mathcal{C}_{\mathrm{f}}(i,m)} m_{n\to m}^{D-(i+1)}(x_{i+1,n}).$$

周辺化に$O\left(|\mathcal{C}_{\mathrm{f}}(i,m)||\mathcal{M}|^{|\mathcal{C}_{\mathrm{f}}(i,m|}\right)$の計算時間がかかる。

The marginalization is computed in $O\left(|\mathcal{C}_{\mathrm{f}}(i,m)||\mathcal{M}|^{|\mathcal{C}_{\mathrm{f}}(i,m|}\right)$ time.

疎なファクターグラフ(Sparse factor graph)

ファクターノードの次数$|\mathcal{C}_{\mathrm{f}}(i,m)|+1$が定数のとき、確率伝播法の計算量は$M$や$N$に関して線形になる。変数ノードも同じ性質を満たすグラフを疎グラフと呼ぶ。

The complexity of BP is linear in M and N when the degrees $|\mathcal{C}_{\mathrm{f}}(i,m)|+1$ of factor nodes are constant. A graph is called sparse if the variable nodes satisfy the same property.

Loopy BP

木でないファクターグラフにも確率伝播法を適用することはできる。

It is possible to apply BP to factor graphs with no tree structure.

ファクターノードy_mから変数ノードx_n へのメッセージ$q^t_{m\to n}(x_n)$

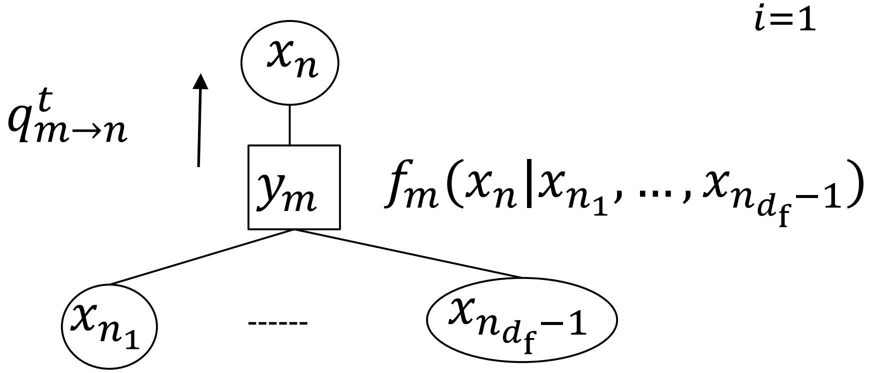
Message $q^t_{m\to n}(x_n)$ passed from the factor node y_m to the variable node x_n

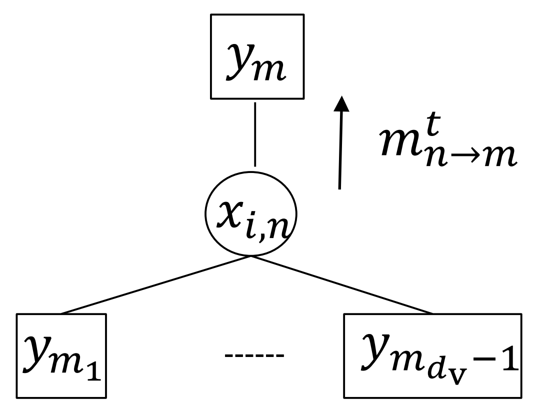
$$q^t_{m\to n}(x_n) = \sum_{\backslash x_n} f_m\left(x_n \middle| x_{n_1}, \dots, x_{n_{d_\mathrm{f}}-1}\right) \prod_{i=1}^{d_\mathrm{f}-1} m^{t-1}_{n_i\to m}(x_{n_i}) .$$

変数ノードx_nからファクターノードy_mへのメッセージ$m^t_{n\to m}(x_n)$

Message $m^t_{n\to m}(x_n)$ passed from the variable node x_n to the variable node y_m

$$m^t_{n\to m}(x_n) \propto p(x_n) \prod_{i=1}^{d_\mathrm{v}-1} q^t_{m_i\to n}(x_n) .$$





Loopy BP

疎なファクターグラフ(Sparse factor graphs)

メッセージの更新を有限回で止める限り、疎グラフは漸近的に木に見える。

A sparse graph looks like a tree asymptotically as long as message updates stop in finite iterations.

- 低密度パリティ検査符号(Low-density parity-check (LDPC) codes)
- ビットインターリーブ符号化変調(Bit-interleaved coded modulation)

密で一様なファクターグラフ(Dense and uniform factor graphs)

BPはサイクルを通じて戻ってくるメッセージを除去する機能を備えている。

BP includes a mechanism that cancels messages returned via cycles.

- Massive multiple-input multiple-output (MIMO)