

2026年3月11日 IEICE総合大会 企画セッション

AI-6：統一的体系化によるAI・データサイエンス研究の新展望

CENTER FOR DATA SCIENCE

データ生成観測メカニズムとしての
決定木とそのもとでの最適予測



中原悠太, 松嶋敏泰 (早稲田大学)

発表概要

- 意思決定写像で従来研究を整理しておくと，“穴”が見えてくる
 - 線形回帰と決定木の対比（“確率モデルの有/無”×“モデルの既知/未知”）
- 意思決定写像で自分の研究を整理しておくと，研究をピボットできる
 - 目的のピボット：予測の不確実性評価
 - 余談：APIとしての意思決定写像
 - 入力のピボット：区間分割モデル
 - 入力のピボット：領域分割モデル
 - 入力のピボット：連続値時系列の文脈木モデル
- モデル未知の設定における計算量的困難の回避

発表概要

- 意思決定写像で従来研究を整理しておくと，“穴”が見えてくる
 - 線形回帰と決定木の対比（“確率モデルの有/無”×“モデルの既知/未知”）
- 意思決定写像で自分の研究を整理しておくと，研究をピボットできる
 - 目的のピボット：予測の不確実性評価
 - 余談：APIとしての意思決定写像
 - 入力のピボット：区間分割モデル
 - 入力のピボット：領域分割モデル
 - 入力のピボット：連続値時系列の文脈木モデル
- モデル未知の設定における計算量的困難の回避

予測問題

- 予測問題とは

説明変数または特徴量

目的変数

性別	年齢	購買額
男 → 0	27	1000
女 → 1	33	2000
⋮	⋮	⋮
男 → 0	18	1500
女 → 1	22	?

カテゴリカルな変数は
ダミー変数化

予測問題

- 予測問題とは

	説明変数または特徴量: x		目的変数: y
	x_1	x_2	y
$\{(x_i, y_i)\}_{i=1}^n$	x_{11}	x_{12}	y_1
	x_{21}	x_{22}	y_2
	\vdots	\vdots	\vdots
	x_{n1}	x_{n2}	y_n
x_{n+1}	$x_{n+1,1}$	$x_{n+1,2}$	$y_{n+1} = ?$

過去の説明変数と目的変数の組 $\{(x_i, y_i)\}_{i=1}^n$ から
新規説明変数 x_{n+1} に対応する未知目的変数 y_{n+1} を予測する

予測問題の意思決定写像

- 予測問題の意思決定写像

目的：新規データの目的変数 y_{n+1} の予測

設定：（未定）

評価基準：（未定）

訓練データ $\{(x_i, y_i)\}_{i=1}^n$
新規データ x_{n+1}

意思決定写像

予測値 \hat{y}_{n+1}

意思決定写像から見た線形回帰

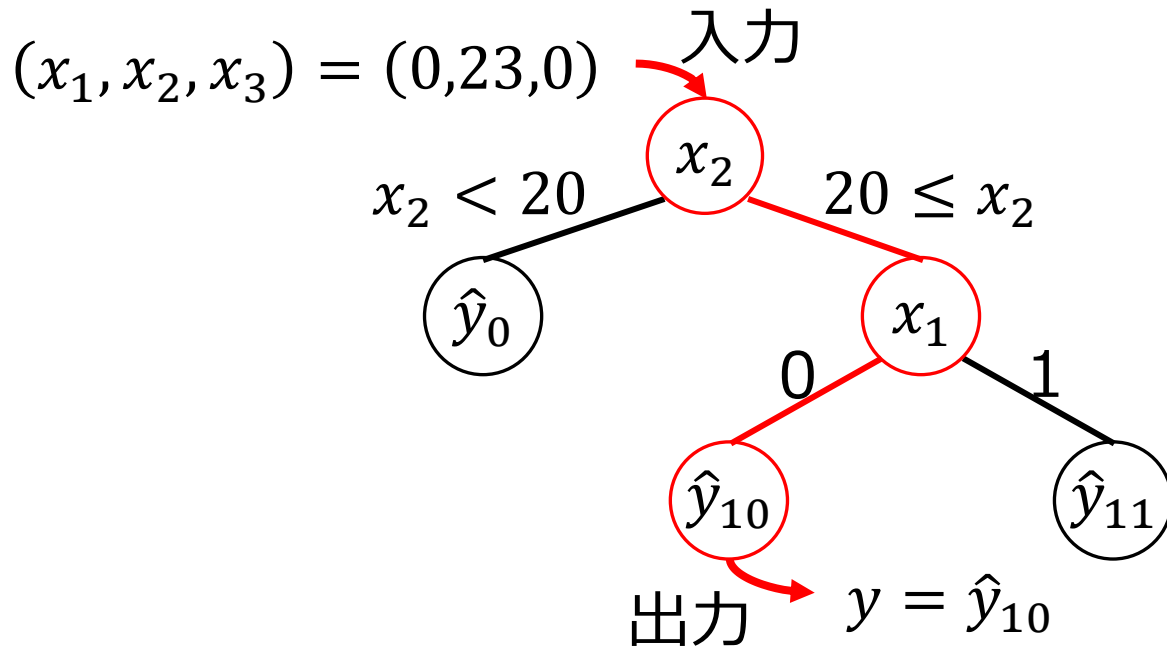
例, モデル m に含まれる変数が x_1, x_3, x_4 のとき
 $\mathbf{x}^{(m)} = [x_1, x_3, x_4]^T$

- 特に, 設定 (データ生成観測メカニズム) の違いに着目

設定の違い	モデル既知	モデル未知
データに同質性を仮定 (確率モデルの仮定なし)	<ul style="list-style-type: none"> 線形関数を用いて予測する $f(x; \boldsymbol{\beta}^{(m)}, m) = \boldsymbol{\beta}^{(m)\top} \mathbf{x}^{(m)}$ モデルmは既知 <p style="text-align: center;">▼</p> <p style="text-align: center;">最小二乗法など</p>	<ul style="list-style-type: none"> 多項式関数を用いて予測する $f(x; \boldsymbol{\beta}^{(m)}, m) = \boldsymbol{\beta}^{(m)\top} \mathbf{x}^{(m)}$ モデルmは未知 <p style="text-align: center;">▼</p> <p style="text-align: center;">リッジ回帰など</p>
データに確率モデルを仮定	<ul style="list-style-type: none"> データは線形回帰モデルに従う $y = \boldsymbol{\beta}^{(m)\top} \mathbf{x}^{(m)} + \varepsilon,$ $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ モデルmは既知 <p style="text-align: center;">▼</p> <p style="text-align: center;">最尤推定量のプラグイン ベイズ最適な直接予測など</p>	<ul style="list-style-type: none"> データは線形回帰モデルに従う $y = \boldsymbol{\beta}^{(m)\top} \mathbf{x}^{(m)} + \varepsilon,$ $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ モデルmは既知であり, 事前分布$p(m)$に従う <p style="text-align: center;">▼</p> <p style="text-align: center;">AICに基づくモデル選択 ベイズ最適な直接予測など</p>

意思決定写像から見た決定木

- 決定木関数 $y = f(\mathbf{x}; \hat{\mathbf{y}}, \mathbf{k}, T)$ (例, CART[Breiman et al., 1984])



この関数を特徴づけるのは. . .

- $\mathbf{k} \in \mathcal{K}$: 各内部ノードに割り当てられた説明変数の添字列 (特徴量割当ベクトル)
- $T \in \mathcal{T}$: 木の形
- $\hat{\mathbf{y}} \in \mathcal{Y}$: 葉ノードに割り当てられた予測値

モデル = (\mathbf{k}, T)

意思決定写像から見た決定木

●モデル既知の同質性を仮定した予測の意思決定写像

目的：新規データの目的変数 y_{n+1} の予測

設定：
・データの同質性を仮定する
・決定木関数 $f(x; \hat{y}, k, T)$ を訓練データに基づいて予め構築しておき、それに新規データの説明変数 x_{n+1} を代入して予測を行う
・モデル k, T は既知

評価基準：訓練データに対する二乗誤差



意思決定写像から見た決定木

●モデル**未知**の同質性を仮定した予測の意思決定写像

目的：新規データの目的変数 y_{n+1} の予測

設定：・データの同質性を仮定する

・決定木関数 $f(x; \hat{y}, k, T)$ を訓練データに基づいて予め構築しておき、
それに新規データの説明変数 x_{n+1} を代入して予測を行う

・モデル k, T は**未知**

評価基準：訓練データに対する二乗誤差

訓練データ $\{(x_i, y_i)\}_{i=1}^n$
新規データ x_{n+1}

意思決定写像

予測値 \hat{y}_{n+1}

意思決定写像から見た決定木

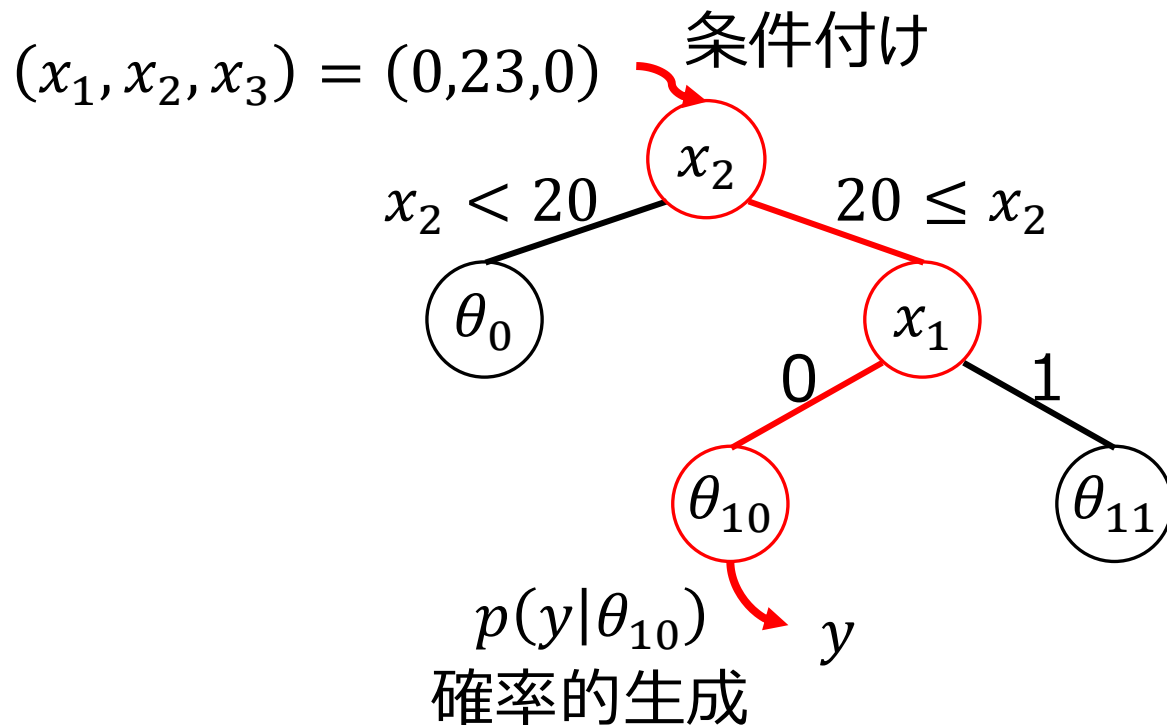
- 特に、設定（データ生成観測メカニズム）の違いに着目

設定の違い	モデル既知	モデル未知
データに同質性を仮定 (確率モデルの仮定なし)	<ul style="list-style-type: none">決定木関数$f(x; \hat{y}, k, T)$を用いて予測するモデルk, Tは既知 <p style="text-align: center;">▼</p> <p style="text-align: center;">層別 + 算術平均など</p>	<ul style="list-style-type: none">決定木関数$f(x; \hat{y}, k, T)$を用いて予測するモデルk, Tは未知 <p style="text-align: center;">▼</p> <p style="text-align: center;">CART [Breiman et al., 1984]など</p>
データに確率モデルを仮定	対応する研究はあるか？	対応する研究はあるか？

意思決定写像から見た決定木

- 決定木確率モデル $p(y|x, \theta, k, T)$

(例, [Chipman et al., 1997], [須子ら, 2003])



この分布を特徴づけるのは. . .

- $k \in \mathcal{K}$: 各内部ノードに割り当てられた説明変数の添字列 (特徴量割当ベクトル)
- $T \in \mathcal{T}$: 木の形
- $\theta \in \Theta$: 葉ノードに割り当てられた生成分布パラメータ

モデル = (k, T)

意思決定写像から見た決定木

●モデル既知の確率モデルを仮定した予測の意思決定写像

目的：新規データの目的変数 y_{n+1} の予測

設定：
・データが決定木確率モデル $p(y|x, \theta, k, T)$ に従って生成，観測される
・パラメータ θ は未知であり，事前分布 $p(\theta)$ に従う
・モデル k, T は既知

評価基準：予測に対する二乗誤差損失に基づくベイズリスク関数



意思決定写像から見た決定木

●モデル未知の確率モデルを仮定した予測の意思決定写像

目的：新規データの目的変数 y_{n+1} の予測

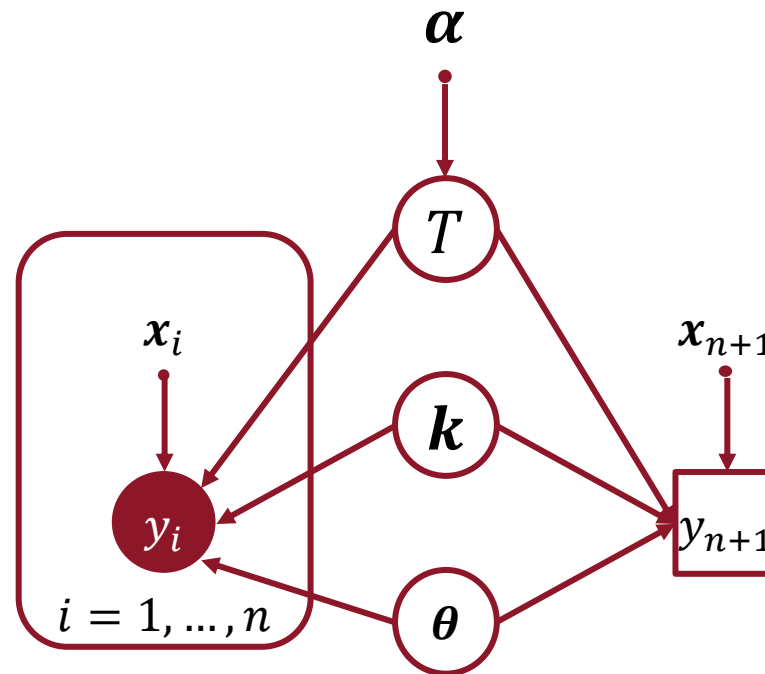
設定：
・データが決定木確率モデル $p(y|x, \theta, k, T)$ に従って生成，観測される
・パラメータ θ は未知であり，事前分布 $p(\theta)$ に従う
・モデル k, T は**未知**であり，事前分布 $p(k, T)$ に従う

評価基準：予測に対する二乗誤差損失に基づくベイズリスク関数



意思決定写像から見た決定木

- 決定木確率モデルのデータ生成観測メカニズムグラフ



意思決定写像から見た決定木

- 特に、設定（データ生成観測メカニズム）の違いに着目

設定の違い	モデル既知	モデル未知
データに同質性を仮定 (確率モデルの仮定なし)	<ul style="list-style-type: none">決定木関数$f(x; \hat{y}, k, T)$を用いて予測するモデルk, Tは既知 <p style="text-align: center;">▼</p> <p style="text-align: center;">層別 + 算術平均など</p>	<ul style="list-style-type: none">決定木関数$f(x; \hat{y}, k, T)$を用いて予測するモデルk, Tは未知 <p style="text-align: center;">▼</p> <p style="text-align: center;">CART [Breiman et al., 1984]など</p>
データに確率モデルを仮定	<ul style="list-style-type: none">データが決定木確率モデル$p(y x, \theta, k, T)$に従って生成, 観測されるモデルk, Tは既知 <p style="text-align: center;">▼</p> <p style="text-align: center;">条件付き正規分布など</p>	<ul style="list-style-type: none">データが決定木確率モデル$p(y x, \theta, k, T)$に従って生成, 観測されるモデルk, Tは未知であり, 事前分布$p(k, T)$に従う <p style="text-align: center;">▼</p> <p style="text-align: center;">メタツリー [Nakahara et al., 2025] BART [Chipman et al., 2010]など</p>

意思決定写像から見た決定木

- 特に, 設定 (データ生成観測メカニズム) の違いに着目

設定の違い	モデル既知	モデル未知
データに同質性を仮定 (確率モデルの仮定)	<ul style="list-style-type: none"> 決定木関数 $f(x; \hat{y}, k, T)$ を用 	<ul style="list-style-type: none"> 決定木関数 $f(x; \hat{y}, k, T)$ を用 予測する モデル k, T は未知
<p>“穴”があったのには理由がある。 最適な予測値の計算が計算量的に困難 →本日の最後にアイデアだけ紹介</p>		
データに確率モデルを仮定	<ul style="list-style-type: none"> データが決定木確率モデル $p(y x, \theta, k, T)$ に従って生成, 観測される モデル k, T は既知 	<ul style="list-style-type: none"> データが決定木確率モデル $p(y x, \theta, k, T)$ に従って生成, 観測される モデル k, T は未知であり, 事前分布 $p(k, T)$ に従う
	<p>条件付き正規分布など</p>	<p>メタツリー [Nakahara et al., 2025] BART [Chipman et al., 2010] など</p>

発表概要

- 意思決定写像で従来研究を整理しておくと，“穴”が見えてくる
 - 線形回帰と決定木の対比（“確率モデルの有/無”×“モデルの既知/未知”）
- 意思決定写像で自分の研究を整理しておくと，研究をピボットできる
 - 目的のピボット：予測の不確実性評価
 - 余談：APIとしての意思決定写像
 - 入力のピボット：区間分割モデル
 - 入力のピボット：領域分割モデル
 - 入力のピボット：連続値時系列の文脈木モデル
- モデル未知の設定における計算量的困難の回避

目的のピボット：予測の不確実性評価 [Nakahara et al., 2025]

● 決定木確率モデルに基づく予測の不確実性評価の意思決定写像

目的：新規データの目的変数 y_{n+1} の予測の不確実性評価

設定：
・データが決定木確率モデル $p(y|x, \theta, \mathbf{k}, T)$ に従って生成，観測される
・パラメータ θ は未知であり，事前分布 $p(\theta)$ に従う
・モデル \mathbf{k}, T は未知であり，事前分布 $p(\mathbf{k}, T)$ に従う

評価基準：
・予測区間 (l, u) による y_{n+1} の被覆確率に基づく事後期待損失
・予測区間 (l, u) の長さ

訓練データ $\{(x_i, y_i)\}_{i=1}^n$
新規データ x_{n+1}

意思決定写像

予測区間 (l, u)

目的のピボット：予測の不確実性評価 [Nakahara et al., 2025]

● 95%予測区間実験結果

- 上段：実際に予測区間に予測値が含まれた頻度 (0.95に近いほど良い)
- 下段：予測区間の長さ (小さいほど良い)
- データセット毎に予測区間が頻度が0.95に近いものを赤字に

Datasets	LightGBM	BART default	BART ntree=1	MTMCMC Leaf model: Normal	MTMCMC Leaf model: LR model
automobile	0.911	0.706	0.433	0.992	0.938
	3.111	1.234	0.926	3.263	3.163
servo	0.908	0.699	0.558	0.990	0.992
	1.938	0.934	0.662	2.204	2.215
cpu	0.935	0.672	0.428	0.983	0.975
	3.888	0.586	0.427	2.028	1.650
liver	0.917	0.282	0.128	0.950	0.943
	3.270	0.833	0.522	3.632	3.703
Mpg	0.854	0.765	0.474	0.979	0.955
	1.379	0.738	0.643	2.181	1.562
student	0.865	0.522	0.189	0.967	0.950
	3.098	1.475	0.500	3.784	3.643
diabetes	0.832	0.576	0.251	0.970	0.957
	2.303	1.211	0.679	3.251	2.832
boston	0.882	0.852	0.331	0.980	0.965
	3.222	0.864	0.501	2.270	1.631
cps	0.845	0.572	0.211	0.954	0.943
	2.696	1.131	0.507	3.415	3.329
strikes	0.820	0.851	0.198	0.953	0.949
	2.012	1.289	0.344	2.608	2.551

余談：APIとしての意思決定写像

- 決定木確率モデルはPythonライブラリBayesMLで公開中 [Nakahara et al., 2026]
- ライブラリのインストール方法：pip install bayesml
- 決定木確率モデルの利用方法
(詳しくはhttps://bayesml.github.io/BayesML/examples/metatree_prediction_interval.html)

```
# パッケージのインポート
from bayesml import metatree

# モデルインスタンスの生成
model = metatree.LearnModel(c_dim_continuous, c_dim_categorical, c_max_depth)

# 事後分布, 予測分布の更新
model.update_posterior(x_train_continuous, x_train_categorical, y_train)
model.calc_pred_dist(x_test_continuous, x_test_categorical)

# 予測値の計算
y_pred = model.make_prediction()
```

余談：APIとしての意思決定写像

- 様々な確率モデルが同じAPIで使える

- ベルヌーイモデル
- カテゴリカルモデル
- ポアソンモデル
- 正規モデル
- 多変量正規モデル
- 指数モデル
- 混合正規モデル
- 線形回帰モデル
- ロジスティック回帰モデル
- 混合線形回帰モデル
- メタツリーモデル
- 自己回帰モデル
- 隠れマルコフモデル
- 文脈木モデル

余談：APIとしての意思決定写像

- Scikit-learnなどで用いられるfit-predict型API
 - 意思決定写像の目的は予測が前提となっているように思える
 - 意思決定写像の評価基準はAPIとしては明示されていない
 - (ベイズ統計, ベイズ機械学習への対応が不十分)

余談：APIとしての意思決定写像

● BayesMLのAPIと意思決定写像の対応

```
# 仮定する確率モデルをインポートし、モデルインスタンスを生成
from bayesml import linearregression
model = linearregression.LearnModel(c_degree=2)
💡
# データを与えて事後分布、予測分布の更新
model.update_posterior(x_train, y_train)
model.calc_pred_dist(x_test)

# 目的に応じたメソッドで、評価基準をオプションに指定し、出力を計算
y_pred = model.make_prediction(loss='squared') # 二乗誤差損失を最小化する予測値を計算
params = model.estimate_params(loss='0-1') # 0-1損失を最小化するパラメータを計算
```

設定を決める

入力を決める

出力を得る

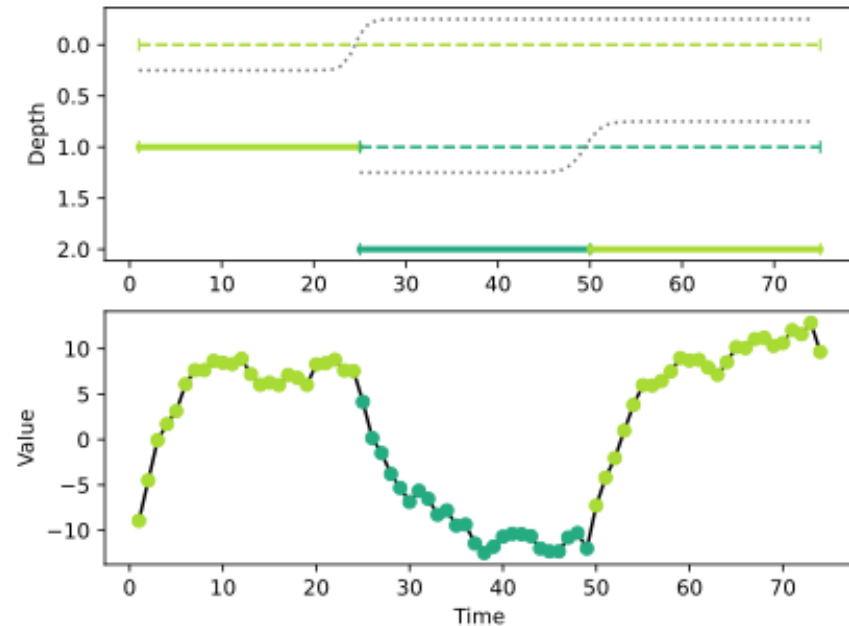
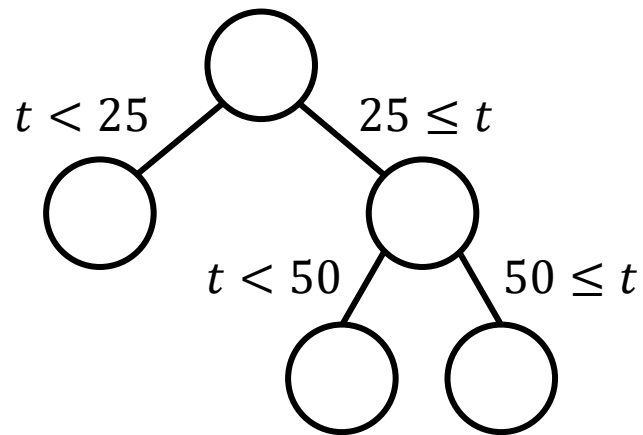
目的を決める

評価基準を決める

面白いと思っていただけの方はぜひ[GitHub](#)上でStarをお願いします！

入力のピボット：区間分割モデル [Nakahara et al., 2026]

- 区間分割モデル



説明変数 x を, 時点 t に

入力のピボット：区間分割モデル [Nakahara et al., 2026]

● 区間分割モデルの意思決定写像

目的：変化点検出

設定：・時点 t のデータ y_t が区間分割モデル $p(y_t|t, \theta, k, T)$ に従って生成,
観測される

・パラメータ θ は未知であり, 事前分布 $p(\theta)$ に従う

・モデル T は未知であり, 事前分布 $p(T)$ に従う

・モデル k は既知

評価基準：0-1損失に基づくベイズリスク関数

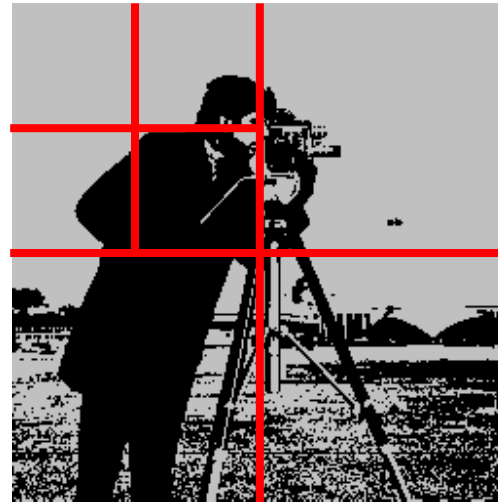
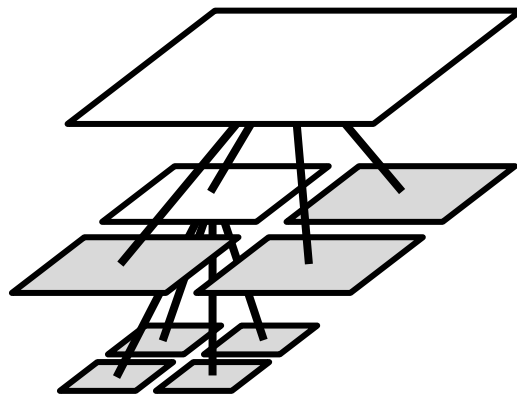
時点とデータの組
 $\{(t, y_t)\}_{t=1}^n$



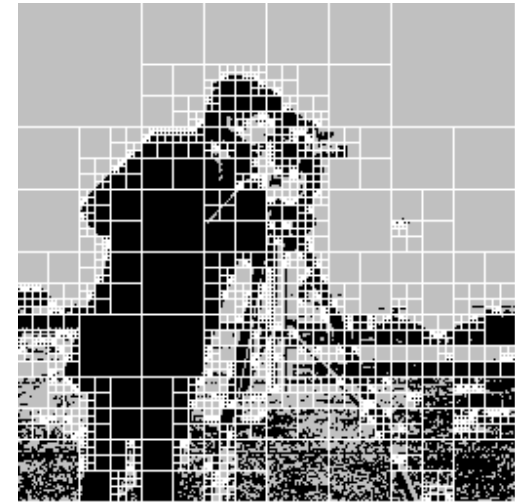
区間分割を表す木
 T

入力のピボット：領域分割モデル [Nakahara et al., 2021]

- 領域分割モデル



推定結果



説明変数 x を、位置インデックス (i, j) に

入力のピボット：領域分割モデル [Nakahara et al., 2021]

● 領域分割モデルの意思決定写像

目的：領域分割

設定：

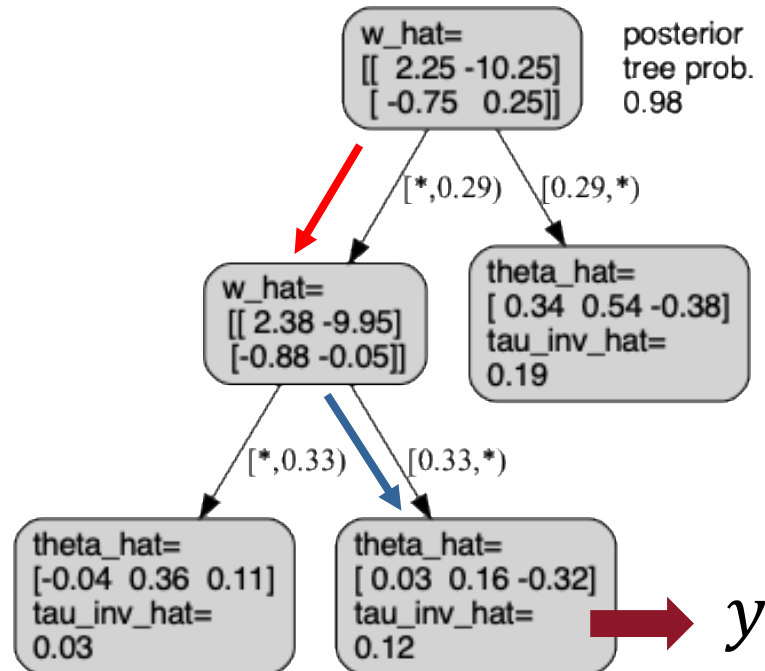
- 位置 (i, j) の画素値 $y_{i,j}$ が領域分割モデル $p(y_{i,j} | i, j, \theta, k, T)$ に従って生成, 観測される
- パラメータ θ は未知であり, 事前分布 $p(\theta)$ に従う
- モデル T は未知であり, 事前分布 $p(T)$ に従う
- モデル k は既知

評価基準：0-1損失に基づくベイズリスク関数



入力のピボット：連続値時系列の文脈木モデル [Saito et al., 2026]

●連続値時系列の文脈木モデル



...0.5, 0.7, 1.1, 0.2, y

深さ d で用いる説明変数 x を, $d + 1$ 時点前のデータの値に

入力のピボット：連続値時系列の文脈木モデル [Saito et al., 2026]

● 連続値時系列の文脈木モデルの意思決定写像

目的：新規データ y_{n+1} の予測

設定：・時点 t のデータ y_t が文脈木モデル $p(y_t|y^{t-1}, \theta, k, T)$ に従って生成,
観測される

・パラメータ θ は未知であり, 事前分布 $p(\theta)$ に従う

・モデル T は未知であり, 事前分布 $p(T)$ に従う

・モデル k は既知

評価基準：予測の二乗誤差損失に基づくベイズリスク関数

過去のデータ $\{y_t\}_{t=1}^n$



意思決定写像



予測値 \hat{y}_{n+1}

発表概要

- 意思決定写像で従来研究を整理しておくと, “穴”が見えてくる
 - 線形回帰と決定木の対比 (“確率モデルの有/無” × “モデルの既知/未知”)
- 意思決定写像で自分の研究を整理しておくと, 研究をピボットできる
 - 目的のピボット: 予測の不確実性評価
 - 余談: APIとしての意思決定写像
 - 入力のピボット: 区間分割モデル
 - 入力のピボット: 領域分割モデル
 - 入力のピボット: 連続値時系列の文脈木モデル
- モデル未知の設定における計算量的困難の回避

モデル未知の設定における計算量的困難性の回避

- モデル未知のもとでのベイズ最適な決定において重要な計算
 - 尤度関数 $p(x|\theta, m)$, パラメータ事前分布 $p(\theta|m)$, モデル事前分布 $p(m)$ に対して,

$$p(m|x) = \frac{p(m) \int p(x|\theta, m)p(\theta|m)d\theta}{\sum_{m \in \mathcal{M}} p(m) \int p(x|\theta, m)p(\theta|m)d\theta}$$

共役事前分布を用いれば計算可能

組合せ爆発

- 対策
 - 厳密性を保った計算量削減 (サブルーチンとしても使える) ← 今回紹介
 - 近似を許した計算量削減
 - モデル探索 / モデル部分集合探索 / MCMC

モデル未知の設定における計算量的困難性の回避

- 本日紹介したモデルの尤度関数の共通点
→ 木 τ の下でのデータ x の分布 $p(x|\tau)$ と書ける (パラメータ θ は省略)
- ベイズ最適な決定のために重要な計算

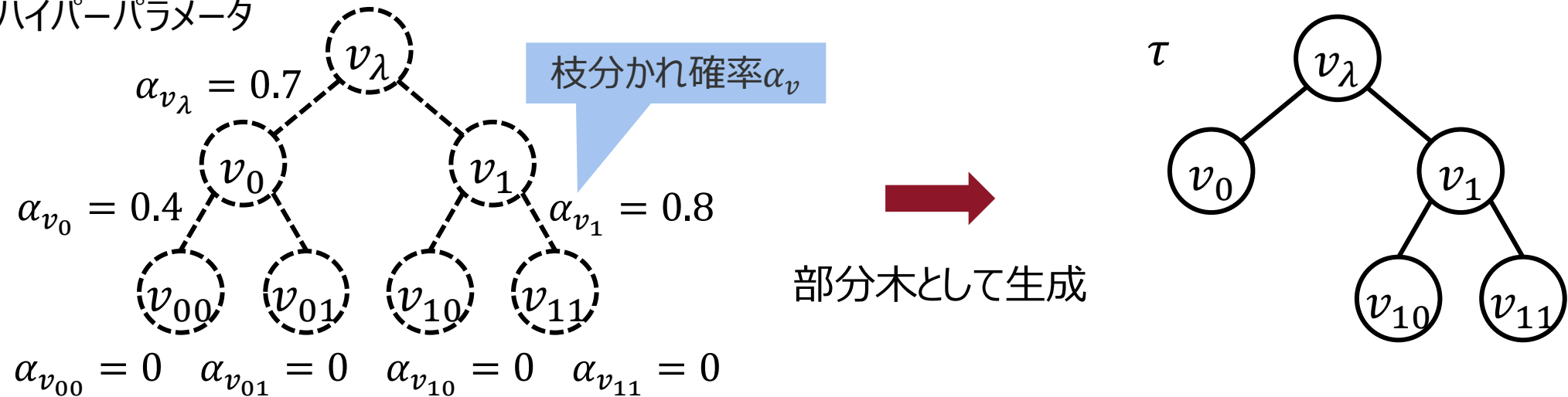
$$p(\tau|x) = \frac{p(x|\tau)p(\tau)}{\sum_{\tau \in \mathcal{T}} p(x|\tau)p(\tau)}$$

この計算が効率的に実行可能となる $p(\tau), p(x|\tau)$ の条件は何か？
条件を満たすとき、どのようなアルゴリズムで計算が可能か？

モデル未知の設定における計算量的困難性の回避

- 事前分布 $p(\tau)$ [Matsushima et al., 2007], [Nakahara et al., 2022]

ハイパーパラメータ



$$\begin{aligned} p(\tau) &= \alpha_{v_\lambda} (1 - \alpha_{v_0}) \alpha_{v_1} (1 - \alpha_{v_{10}}) (1 - \alpha_{v_{11}}) \\ &= 0.7 \times (1 - 0.4) \times 0.8 \times (1 - 0) \times (1 - 0) = 0.336 \end{aligned}$$

$$\text{一般型 : } p(\tau) := \prod_{v \in \mathcal{J}_\tau} \alpha_v \prod_{v' \in \mathcal{L}_\tau} (1 - \alpha_{v'})$$

モデル未知の設定における計算量的困難性の回避

- 定理 [Nakahara et al., 2022]

- 尤度関数の関数型の制約：

内部ノードに関する項と葉ノードに関する項の積の形に因数分解可能

$$p(x|\tau) = \prod_{v \in \mathcal{J}_\tau} g(x, v) \prod_{v' \in \mathcal{L}_\tau} h(x, v')$$

- 事後分布：

$$p(\tau|x) = \prod_{v \in \mathcal{J}_\tau} \alpha_{v|x} \prod_{v' \in \mathcal{L}_\tau} (1 - \alpha_{v'|x})$$

規格化定数も含め、事前分布と同じ形で表現できる（共役事前分布）

モデル未知の設定における計算量的困難性の回避

- 定理 [Nakahara et al., 2022]

- 尤度関数の関数型の制約：

内部ノードに関する項と葉ノードに関する項の積の形に因数分解可能

$$p(x|\tau) = \prod_{v \in \mathcal{J}_\tau} g(x, v) \prod_{v' \in \mathcal{L}_\tau} h(x, v')$$

- 事後分布：

$$p(\tau|x) = \prod_{v \in \mathcal{J}_\tau} \alpha_{v|x} \prod_{v' \in \mathcal{L}_\tau} (1 - \alpha_{v'|x})$$

ノード数のオーダーで計算可能！

$$\alpha_{v|x} := \begin{cases} \alpha_v, & v \in \mathcal{L}_p \\ \frac{\alpha_v g(x, v) \prod_{v' \in \text{Ch}_p} q(x|v')}{q(x|v)}, & v \in \mathcal{J}_p \end{cases} \quad q(x|v) := \begin{cases} h(x, v), & v \in \mathcal{L}_p \\ (1 - \alpha_v) h(x, v) + g(x, v) \prod_{v' \in \text{Ch}_p(v)} q(x|v'), & v \in \mathcal{J}_p \end{cases}$$

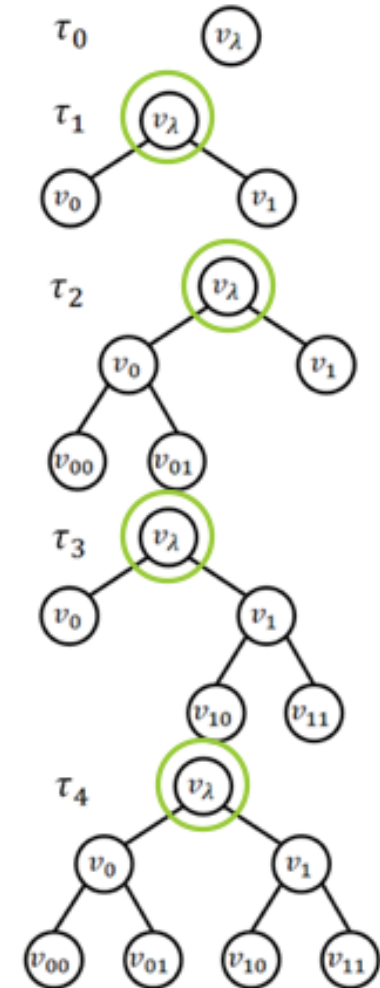
モデル未知の設定における計算量的困難性の回避

- 計算量削減のアイデア

- $p(x|\tau) = 1$ と置いたもとでの $\sum_{\tau \in \mathcal{T}} p(\tau)p(x|\tau)$ の期待値

$$\begin{aligned}
 & \sum_{\tau \in \mathcal{T}} p(\tau) \\
 &= (1 - \alpha_{v_\lambda}) \\
 & \quad + \alpha_{v_\lambda} (1 - \alpha_{v_0})(1 - \alpha_{v_1}) \\
 & \quad + \alpha_{v_\lambda} \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}})(1 - \alpha_{v_1}) \\
 & \quad + \alpha_{v_\lambda} (1 - \alpha_{v_0}) \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}}) \\
 & \quad + \alpha_{v_\lambda} \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}}) \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}})
 \end{aligned} \tag{2}$$

$$\begin{aligned}
 &= (1 - \alpha_{v_\lambda}) \\
 & \quad + \alpha_{v_\lambda} \{ (1 - \alpha_{v_0})(1 - \alpha_{v_1}) \\
 & \quad \quad + \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}})(1 - \alpha_{v_1}) \\
 & \quad \quad + (1 - \alpha_{v_0}) \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}}) \\
 & \quad \quad + \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}}) \\
 & \quad \quad \quad \times \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}}) \}
 \end{aligned} \tag{3}$$



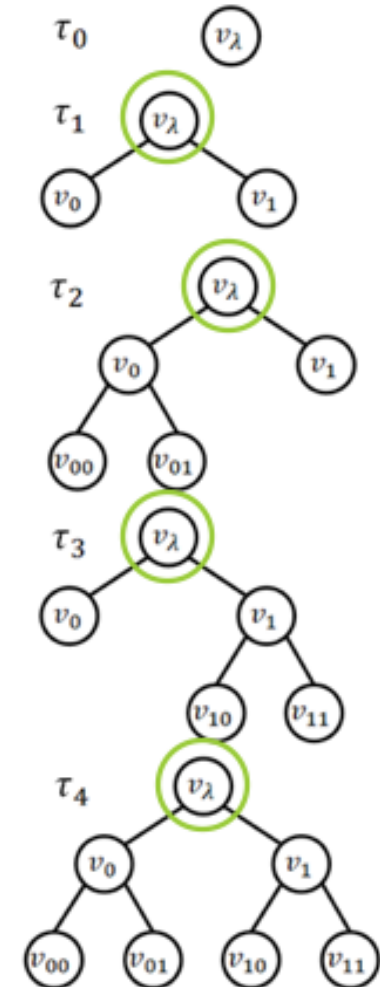
モデル未知の設定における計算量的困難性の回避

- 計算量削減のアイデア

- $p(x|\tau) = 1$ と置いたもとでの $\sum_{\tau \in \mathcal{T}} p(\tau)p(x|\tau)$ の期待値

$$\begin{aligned}
 & \sum_{\tau \in \mathcal{T}} p(\tau) \\
 &= (1 - \alpha_{v_\lambda}) \\
 & \quad + \alpha_{v_\lambda} (1 - \alpha_{v_0})(1 - \alpha_{v_1}) \\
 & \quad + \alpha_{v_\lambda} \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}})(1 - \alpha_{v_1}) \\
 & \quad + \alpha_{v_\lambda} (1 - \alpha_{v_0}) \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}}) \\
 & \quad + \alpha_{v_\lambda} \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}}) \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}})
 \end{aligned} \tag{2}$$

$$\begin{aligned}
 &= (1 - \alpha_{v_\lambda}) \\
 & \quad + \alpha_{v_\lambda} \{ (1 - \alpha_{v_0})(1 - \alpha_{v_1}) \\
 & \quad \quad + \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}})(1 - \alpha_{v_1}) \\
 & \quad \quad + (1 - \alpha_{v_0}) \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}}) \\
 & \quad \quad + \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}}) \\
 & \quad \quad \quad \times \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}}) \}
 \end{aligned} \tag{3}$$



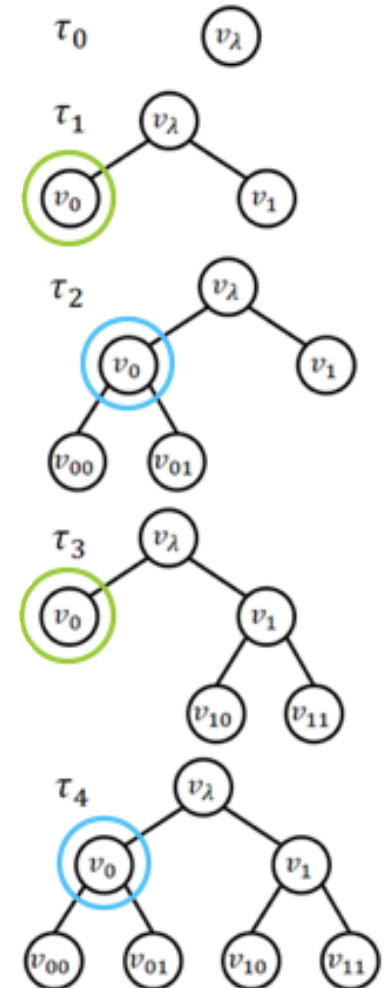
モデル未知の設定における計算量的困難性の回避

- 計算量削減のアイデア

- $p(x|\tau) = 1$ と置いたもとでの $\sum_{\tau \in \mathcal{T}} p(\tau)p(x|\tau)$ の期待値

$$\begin{aligned}
 &= (1 - \alpha_{v_\lambda}) \\
 &\quad + \alpha_{v_\lambda} \left\{ \frac{(1 - \alpha_{v_0}) [(1 - \alpha_{v_1}) + \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}})]}{\frac{\alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}})}{\times [(1 - \alpha_{v_1}) + \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}})]}} \right\} \\
 & \hspace{15em} (4)
 \end{aligned}$$

$$\begin{aligned}
 &= (1 - \alpha_{v_\lambda}) \\
 &\quad + \alpha_{v_\lambda} \left\{ [(1 - \alpha_{v_0}) + \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}})] \right. \\
 &\quad \left. \times [(1 - \alpha_{v_1}) + \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}})] \right\}. \\
 & \hspace{15em} (5)
 \end{aligned}$$



モデル未知の設定における計算量的困難性の回避

- 計算量削減のアイデア

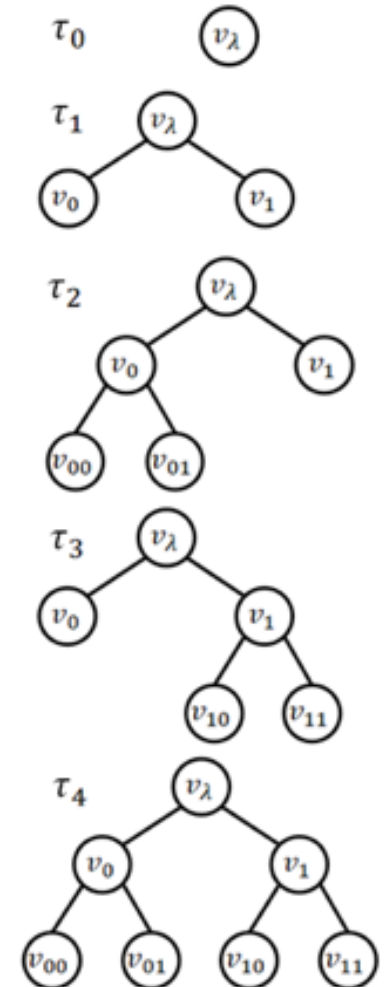
- $p(x|\tau) = 1$ と置いたもとでの $\sum_{\tau \in \mathcal{T}} p(\tau)p(x|\tau)$ の期待値

$$\begin{aligned}
 &= (1 - \alpha_{v_\lambda}) \\
 &\quad + \alpha_{v_\lambda} \left\{ (1 - \alpha_{v_0}) \left[(1 - \alpha_{v_1}) + \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}}) \right] \right. \\
 &\quad\quad \left. + \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}}) \right. \\
 &\quad\quad \left. \times \left[(1 - \alpha_{v_1}) + \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}}) \right] \right\} \tag{4}
 \end{aligned}$$

$$\begin{aligned}
 &= (1 - \alpha_{v_\lambda}) \\
 &\quad + \alpha_{v_\lambda} \left\{ \left[(1 - \alpha_{v_0}) + \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}}) \right] \right. \\
 &\quad\quad \left. \times \left[(1 - \alpha_{v_1}) + \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}}) \right] \right\}. \tag{5}
 \end{aligned}$$

因数分解のアイデア

左止・右止 + 左分・右止 + 左止・右分 + 左分・右分
 = (左分 + 左止) · (右分 + 右止)



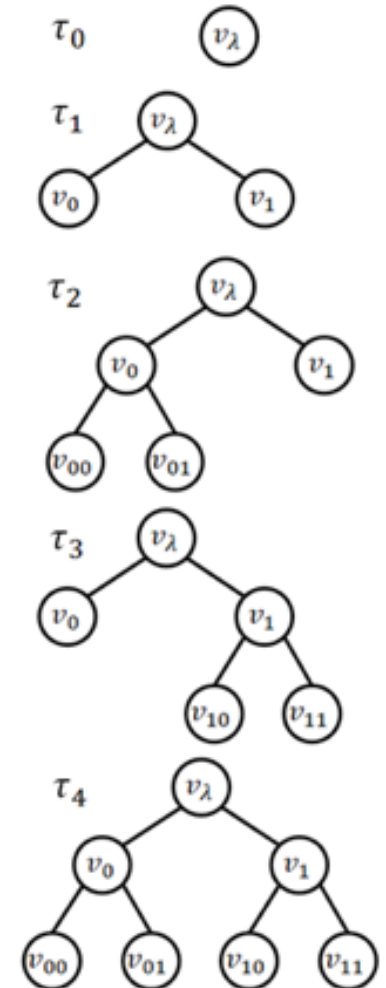
モデル未知の設定における計算量的困難性の回避

- 計算量削減のアイデア

- $p(x|\tau) = 1$ と置いたもとでの $\sum_{\tau \in \mathcal{T}} p(\tau)p(x|\tau)$ の期待値

$$\begin{aligned}
 &= (1 - \alpha_{v_\lambda}) \\
 &\quad + \alpha_{v_\lambda} \left\{ (1 - \alpha_{v_0}) [(1 - \alpha_{v_1}) + \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}})] \right. \\
 &\quad\quad \left. + \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}}) \right. \\
 &\quad\quad \left. \times [(1 - \alpha_{v_1}) + \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}})] \right\} \\
 &\hspace{15em} (4)
 \end{aligned}$$

$$\begin{aligned}
 &= (1 - \alpha_{v_\lambda}) \\
 &\quad + \alpha_{v_\lambda} \left\{ [(1 - \alpha_{v_0}) + \alpha_{v_0} (1 - \alpha_{v_{00}})(1 - \alpha_{v_{01}})] \right. \\
 &\quad\quad \left. \times [(1 - \alpha_{v_1}) + \alpha_{v_1} (1 - \alpha_{v_{10}})(1 - \alpha_{v_{11}})] \right\}. \\
 &\hspace{10em} \text{再帰的構造 } (1 - \alpha) + \alpha [***] \hspace{10em} (5)
 \end{aligned}$$



まとめ

- 意思決定写像で従来研究を整理しておくと，“穴”が見えてくる
- 意思決定写像で自分の研究を整理しておくと，研究をピボットできる
- 尤度関数がノードに関する積で書けていれば，枝分かれ確率の積で表される木の事前分布を活用して計算量的困難を回避できる

参考文献

- L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, Classification and regression trees. CRC press, 1984.
- Hugh Chipman, Edward I. George, Robert E. McCulloch, A Bayesian approach to CART, Proceedings of the Sixth International Workshop on Artificial Intelligence and Statistics, PMLR R1:91-102, 1997.
- 須子統太, 野村亮, 松嶋敏泰, 平澤茂一, 決定木モデルにおける予測アルゴリズムについて, 電子情報通信学会技術研究報告, vol. 103, np. 246, pp.93-98, 2003.
- Hugh A. Chipman, Edward I. George, and Robert E. McCulloch. BART: Bayesian additive regression trees. The Annals of Applied Statistics, 4(1):266 – 298, 2010.
- Yuta Nakahara, Shota Saito, Naoki Ichijo, Koki Kazama, Toshiyasu Matsushima, Bayesian Decision Theory on Decision Trees: Uncertainty Evaluation and Interpretability, Proceedings of The 28th International Conference on Artificial Intelligence and Statistics, PMLR 258:1045-1053, 2025.
- Y. Nakahara, N. Ichijo, K. Shimada, Y. Iikubo, S. Saito, K. Kazama, T. Matsushima, BayesML Developers, BayesML, Python package version 0.4.1, 2026. [Online] <https://github.com/bayesml/BayesML>

参考文献

- Nakahara, Y., Saito, S., Horinouchi, K., Shimada, K., Ichijo, N., Kobayashi, M., & Matsushima, T. (2026). Variable Splitting Binary Tree Models Based on Bayesian Context Tree Models for Time Series Segmentation. arXiv preprint arXiv:2601.16112.
- Nakahara, Y.; Matsushima, T. A Stochastic Model for Block Segmentation of Images Based on the Quadtree and the Bayes Code for It. *Entropy* **2021**, *23*, 991. <https://doi.org/10.3390/e23080991>
- Saito, S., Nakahara, Y., & Matsushima, T. (2026). Soft Bayesian Context Tree Models for Real-Valued Time Series. arXiv preprint arXiv:2601.11079.
- T. Matsushima and S. Hirasawa, "A Class of Prior Distributions on Context Tree Models and an Efficient Algorithm of the Bayes Codes Assuming It," 2007 IEEE International Symposium on Signal Processing and Information Technology, Giza, Egypt, 2007, pp. 938-941, doi: 10.1109/ISSPIT.2007.4458049.
- Nakahara, Y.; Saito, S.; Kamatsuka, A.; Matsushima, T. Probability Distribution on Full Rooted Trees. *Entropy* **2022**, *24*, 328. <https://doi.org/10.3390/e24030328>