最適化理論

—ベイズ計算の基礎 2—

今回学ぶこと

- ２変数の確率推論（復習）
- 多変数の場合の確率推論

モンティ・ホール問題

- **モンティ・ホール が司会のアメリカのゲームショー番組**
  “Let’s make a deal” (60 **年代**) **の中で行われたゲームに由来**

モンティ・ホール問題の詳細

Wikipedia 「モンティ・ホール問題」より引用

- プレイヤーの前に 3 つのドアがあって、1 つのドアの後ろには景品の新車が、2 つのドアの後ろにはヤギ (はずれを意味する) がいる。
- プレイヤーは新車のドアを当てると新車がもらえる。プレイヤーが 1 つのドアを選択した後、モンティが残りのドアのうちヤギがいるドアを開けてヤギを見せる。
- ここでプレイヤーは最初に選んだドアを、残っている開けられていないドアに変更してもよいと言われる。
- プレイヤーはドアを変更すべきだろうか？

(補足：最初の並びかたは完全に一様ランダム。またモンティがヤギのドアを開けるときに２つ選択肢があるときはモンティはコインフリップの結果に従いドアを選ぶ)

答

ドアを変更するほうが得である。変更しない場合と比べて、車の当たる確率が２倍になる。

設定

- ドアに 1,2,3 と番号をつける
- プレイヤーはドア１を選択したものと仮定する
- 確率変数 X は車の位置を表す。$D(X) = \{1, 2, 3\}$

$$P_X(1) = P_X(2) = P_X(3) = 1/3$$

- 確率変数 M はモンティが開けるドアを表す。$D(M) = \{2, 3\}$

最初にドアを選択する段階で車が当たる確率は $1/3$。

同時分布

x	m	$P_{XM}(x, m)$
1	2	$1/6$
1	3	$1/6$
2	3	$1/3$
3	2	$1/3$

注：上記以外の x, m の組み合わせでは、$P_{XM}(x, m) = 0$

$P_{X|M}(3|2)$ **（「モンティが 2 つ目のドアを開けたとき、選択ドアを変更」に対応）を計算してみよう。**

事後分布計算

x	m	$P_{XM}(x,m)$
1	2	$1/6$
1	3	$1/6$
2	3	$1/3$
3	2	$1/3$

$$P_M(2) = 1/6 + 1/3 = 1/2$$

$$P_{X|M}(3|2) = \frac{P_{XM}(3,2)}{P_M(2)} = 2/3$$

ちなみに

$$P_{X|M}(1|2) = 1 - 2/3 = 1/3$$

まとめると

事前確率

$$P_X(1) = 1/3$$

だったのが、ドア選択を変更することにより

$$P_{X|M}(3|2) = 2/3$$

と確率が２倍になる。なお、モンティがドア 3 を開けた場合も同様。

多変数の確率推論問題：準備

周辺化則

$$
\begin{aligned}
P_{XY}(x, y) &= \sum_z P_{XYZ}(x, y, z) \\
P_X(x) &= \sum_y \sum_z P_{XYZ}(x, y, z)
\end{aligned}
$$

多変数の確率推論問題：例

3 変数の同時分布 $P_{X_1X_2X_3}(x_1, x_2, x_3)$

$x_1x_2x_3$	$P_{X_1X_2X_3}(x_1, x_2, x_3)$
000	0.2
001	0.2
010	0.1
011	0.1
100	0.1
101	0.1
110	0.1
111	0.1

- $P_{X_1}(x_1)$ を求めよ。

多変数の確率推論問題：例

$$
\begin{aligned}
P_{X_1}(x_1) &= \sum_{x_2}\sum_{x_3} P_{X_1X_2X_3}(x_1, x_2, x_3) \\
&= P_{X_1X_2X_3}(x_1, 0, 0) + P_{X_1X_2X_3}(x_1, 0, 1) \\
&+ P_{X_1X_2X_3}(x_1, 1, 0) + P_{X_1X_2X_3}(x_1, 1, 1)
\end{aligned}
$$

となるので、

$$
\begin{aligned}
P_{X_1}(0) &= 0.2 + 0.2 + 0.1 + 0.1 = 0.6 \\
P_{X_1}(1) &= 0.1 + 0.1 + 0.1 + 0.1 = 0.4
\end{aligned}
$$

である。

多変数の確率推論問題：例

典型的な確率推論問題

$P_{X_1|X_2}(0|1)$ を計算せよ。

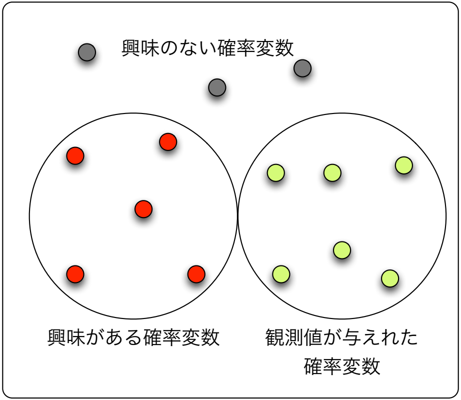
- 系が N 個 $(N = 3)$ の確率変数を含んでいる
- 同時分布が既知 (= 系に関する知識)
- 確率変数の一部の値を観測 $(X_2 = 1)$
- 観測できていない確率変数 (X_1) の事後分布を知りたい
- 興味のない確率変数もある (X_3)

多変数の確率推論問題：例

$$
\begin{aligned}
P_{X_1|X_2}(0|1) &= \frac{P_{X_1X_2}(0,1)}{P_{X_2}(1)} \\
&= \frac{\sum_{x_3} P_{X_1X_2X_3}(0,1,x_3)}{\sum_{x_1}\sum_{x_3} P_{X_1X_2X_3}(x_1,1,x_3)} \\
&\rightarrow \text{your excercise}
\end{aligned}
$$

多変数の確率推論問題の構造

対象とする確率モデルの世界



確率推論の目標：事後確率分布 

多変数の確率推論問題の構造 (2)

- X_1, X_2, X_3, X_4, X_5
- $P_{X_1 \cdots X_5}(x_1, \ldots, x_5)$：**与えられている**
- $V = \{X_3, X_4\}$: **観測変数，**$\{x_3^*, x_4^*\}$：**観測値**
- $I = \{X_1, X_2\}$: **推論対象変数 (興味のある変数)**
- **事後確率計算の目標**

$$P_{X_1 X_2 | X_3 X_4}(x_1 x_2 | x_3^* x_4^*)$$

問：事後確率を同時分布に基づいて書き直せ。

同時分布を使って書く

$$
\begin{aligned}
P_{X_1X_2|X_3X_4}(x_1, x_2|x_3^*, x_4^*) &= \frac{P_{X_1X_2X_3X_4}(x_1, x_2, x_3^*, x_4^*)}{P_{X_3X_4}(x_3^*, x_4^*)} \\
&= \frac{\sum_{x_5} P_{X_1X_2X_3X_4X_5}(x_1, x_2, x_3^*, x_4^*, x_5)}{\sum_{x_1, x_2, x_5} P_{X_1X_2X_3X_4X_5}(x_1, x_2, x_3^*, x_4^*, x_5)}
\end{aligned}
$$

一般の事後確率計算問題も同様にして、周辺化に基づいて「原理的」には計算可能。

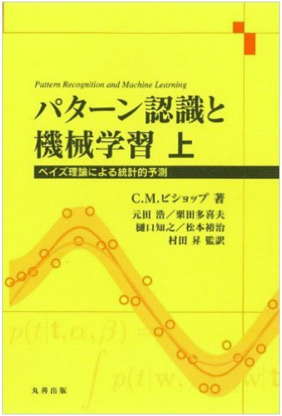
計算量の壁

系に含まれる変数の数を n とする。

- 単純に式の通りに周辺化計算を行う場合、事後確率計算は n に対して指数時間かかる→計算量の壁！
- 音声処理・画像認識・通信などさまざまな現実的な応用において，系に含まれる確率変数は数千〜数万にのぼる場合も多い
- 同時分布を記憶するための「記憶容量の壁」もある！
- 近年、いくつかのブレークスルーがあった
  - グラフィカルモデルの発展
  - マルコフ連鎖モンテカルロ法
  - ビリーフプロパゲーション

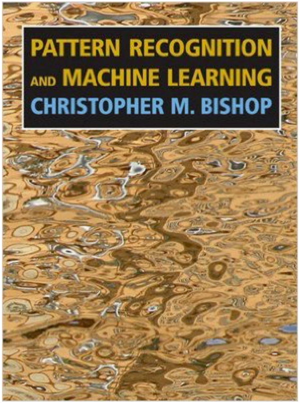
機械学習の教科書 (1)

ここ 10 年ぐらいの定番。この講義の内容は、本書の後半に対応する。



¥7020(上）@amazon

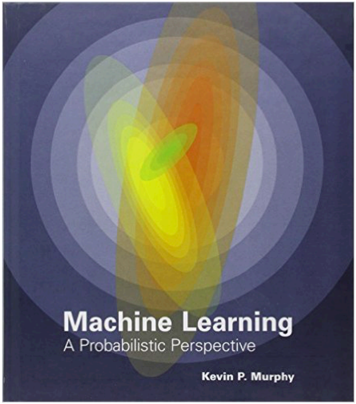
¥8424(下）



¥7999@amazon

機械学習の教科書 (2)

最近に出た本でおすすめはコレ。1000 ページ以上のボリューム。
多彩な話題が網羅されている（ある種、百科事典的)。



¥9791@amazon