α-分布族のファジィ演算子的作用

北陸先端科学技術大学院大学 林 幸雄

Contents

1. α-アファイン多様体への拡張
2. 平均値関数と平均演算子
3. 巾乗関数の性質（α-分布族の平均化作用と問題点）
4. α-分布族によるデータベースモデルの提案
5. 既存の DB, IR モデルに対する位置付け
6. 発見的な情報探索と視点の推定
7. まとめ

==

—

1. α-アファイン多様体への拡張

α-分布族 $S_\alpha = \{p_\alpha(x;\ \theta)\}$,

$$p_\alpha(x;\ \theta) \overset{\text{def}}{=} \left[\frac{1-\alpha}{2}\sum_{i=1}^{n+1}\theta^i F_i(x)\right]^{\frac{2}{1-\alpha}},\quad (\alpha \neq 1)$$

$\mathcal{X} = \{1, 2, \ldots, m\}$ 上の確率分布 $p_i(x) > 0$, $F_i(x) \overset{\text{def}}{=} [p_i(x)]^{\frac{1-\alpha}{2}}$.

$\Downarrow$ α-アファイン多様体$\tilde{S}$に拡張

$\tilde{S} = \{\tilde{p}(x;\ \xi) |\ \tilde{p}(x;\ \xi) > 0, \sum_{x\in\mathcal{X}} \tilde{p}(x;\ \xi) < \infty\}$,

$$\tilde{p}(x;\ \xi) \overset{\text{def}}{=} \left[\frac{1}{n}\sum_{i=1}^{n}\xi^i F_i(x)\right]^{\frac{2}{1-\alpha}} = \left[\frac{1}{n}\sum_{i=1}^{n}u_i(x)^\beta\right]^{\frac{1}{\beta}},$$

但し, $\beta \overset{\text{def}}{=} \frac{1-\alpha}{2}$, $\xi^i \overset{\text{def}}{=} n\beta\theta^i$, $u_i(x) \overset{\text{def}}{=} (\xi^i)^{\frac{1}{\beta}} p_i(x)$ とする.
$\beta > 1$ なら l^β-ノルム, $\beta < 1$ なら Constant Elasticity of Substitution 関数である.

==

—

2. 平均値関数と平均演算子

2.1 平均値関数（オペレーションズリサーチ）

定義 1 連続関数 $z = \phi(u_1, u_2)$ が次の *3* つの条件を満足するとき, これを平均値関数, その値を平均値という.

1. $u_1, u_2 > 0$ で, $min(u_1, u_2) \le \phi(u_1, u_2) \le max(u_1, u_2)$.
2. 関数$\phi(u_1, u_2)$ は u_1, u_2のそれぞれについて非減少関数であり, 停留点がない.
   $\phi(u_1, u_2)$ が微分可能な場合: $grad\ \phi \ge 0$, $div\ \phi > 0$.
3. 関数$\phi(u_1, u_2) = \phi(u_1, u_2)$ は *1* 次の同次関数である. すなわち, すべての$\gamma > 0$ について, $\phi(\gamma u_1, \gamma u_2) = \gamma\phi(u_1, u_2)$ が成立する（尺度変換による不変性の一種）.

平均： min $\le$ 調和 $\le$ 幾何 $\le$ 算術 $\le$双対幾何$\le$ 双対調和 $\le$ max

==

—

2. 平均値関数と平均演算子

2.2 平均演算子（ファジィ論理演算）

定義 2 平均演算子 M は, M: $[0,1] \times [0,1] \to [0,1]$ なる関数であり, 次の条件を満たすものである.

1. $min(u_1, u_2) \leq M(u_1, u_2) \leq max(u_1, u_2)$.
2. $M(u_1, u_2) = M(u_2, u_1)$ （交換性）.
3. M は連続, 増加関数である.
4. $M(0,0) = 0$; $M(1,1) = 1$.
5. $M(u,u) = u$（巾等性）.
6. $M \notin \{min, max\}$.

ファジィ演算： t-ノルム ≤ min ≤ 平均演算子≤ max ≤ t-コノルム

===

—

2. 平均値関数と平均演算子

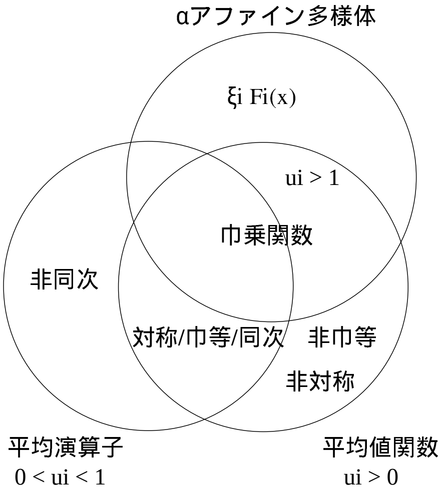

αアファイン多様体
ξi Fi(x)
ui > 1
巾乗関数
非同次
対称/巾等/同次
非巾等
非対称
平均演算子
0 < ui < 1
平均値関数
ui > 0

図 1: それぞれの関数クラスの関係

2. 平均値関数と平均演算子

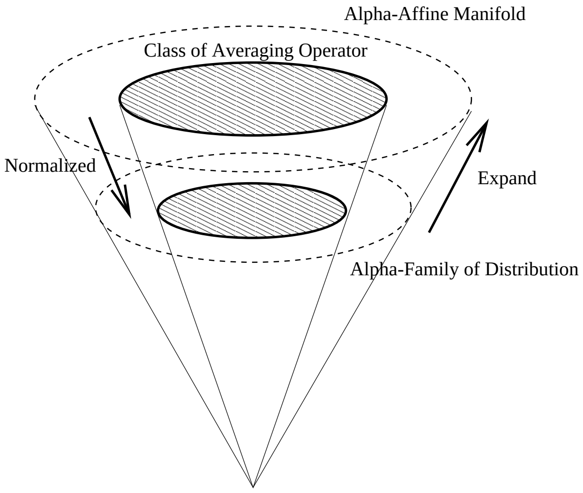

Alpha-Affine Manifold
Class of Averaging Operator
Normalized
Expand
Alpha-Family of Distribution

図 2: α-分布族と平均演算子のクラス（ $0 \leq (\xi^i)^{\frac{1}{\beta}} p_i(x) \leq 1$ に制限 ）

3. 巾乗関数の性質

3.1 α-分布族の平均化作用

β	α	α-接続	平均の解釈	AND/OR 的作用
$-\infty$	∞	-	最小	AND
0	1	指数型	幾何平均	AND 的
1/2	0	リーマン	平方根平均	AND 的
1	-1	混合型	算術平均	中庸
2	-3	-	遠心力平均	OR 的
∞	$-\infty$	-	最大	OR

表 1: α-分布族と平均演算子の物理的解釈の対応

例：幾何平均$z = [\, \Pi_i u_i \,]^{\frac{1}{n}}$は, 両辺の$\log$をとると, 指数型$\log z = \frac{1}{n}\sum_i \log u_i$ となる（$\beta \to 0$の場合）.

==

—

3. 巾乗関数の性質

3.2 $\alpha \neq \pm 1$ 場合の問題点

正規化条件より, ξ^{n+1}と $F_{n+1}(x)$ を決めるが（ $x = 1, \ldots, m$ ）, これらは$n \times m$ 個のデータ $p_i(x)$ （あるいは $F_i(x) = p_i(x)^\beta$ ）に依存する反面, 他制約がなく自由度が大. 仮に, 一様分布 $p_{n+1}(x) = \frac{1}{m}$ として, 1 変数ξ^{n+1}のみを決めると, 正規化条件

$$\sum_{x \in \mathcal{X}} p(x;\ \xi) = \beta^{\frac{1}{\beta}} \sum_{x \in \mathcal{X}} \left[\sum_{i=1}^{n} \xi^i F_i(x) + \frac{\xi^{n+1}}{m^\beta} \right]^{\frac{1}{\beta}} = 1,$$

より, 今度は代数方程式

$$\sum_{x \in \mathcal{X}} (\xi^{n+1} + d_x)^{\frac{1}{\beta}} = m\beta^{-\frac{1}{\beta}},$$

を解かなくてはならなくなる $(d_x \overset{\text{def}}{=} m^\beta \times \sum_{i=1}^{n} \xi^i F_i(x))$.

==

4. α-分布族によるデータベースモデルの提案

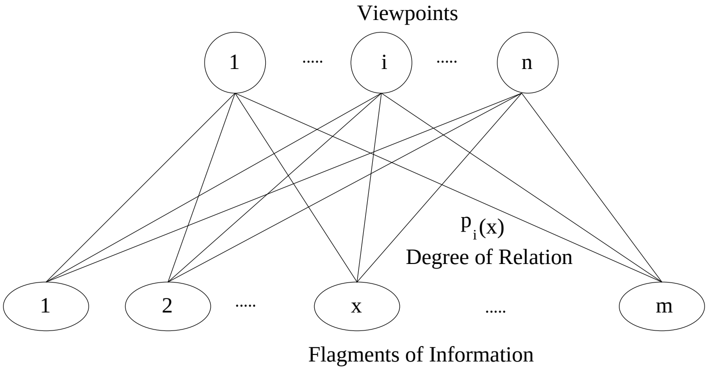

Viewpoints
1
.....
i
.....
n
p_i(x)
Degree of Relation
1
2
.....
x
.....
m
Flagments of Information

図 3: 関係 (確率) 分布モデル

⇒ 視点の重要度パラメータθで張られる情報の関連度の分布空間として多様体 $S = \{p(x;\ \theta)\}$ を考え，Sに対する操作 (パラメータ変化による空間探索) を，探索経路上の視点に応じて $\boxed{\text{情報の関係構造を動的に生成する DB 操作}}$ とする．

—

すなわち, 関係分布モデルでは, $\boxed{10}$

- 予め規定された複数 (n 個) の視点に対して, 利用者が各視点の重要度パラメータ$\theta = (\theta^1, \theta^2, \ldots, \theta^n)$ の値をモデルに与えることで, 利用者の任意の視点を表現する.
- 各視点 i $(= 1, 2, \ldots, n)$ に対して予め規定された, 情報 $x \in \mathcal{X}$ $(= 1, 2, \ldots, m)$ の関係の強弱の度合を表す確率分布 $p_i(x) > 0$ を, 視点の重要度パラメータθで重み付けて, 利用者の視点に対する情報の関連度として, 確率分布 $p(x;\ \theta)$ を生成する.

ことを考える.

===

5. 既存の DB, IR モデルに対する位置付け

5.1 DB モデルの関係の操作性

< DB の関係操作の発展 >

RDB : 関係代数による関係表の分割統合,

ER : 関係属性の付与 (RDB 設計に便利) ,

SDM / OODB : 関係の階層 / 抽象化,

HyperText DB: リンク編集,

関係分布モデル : 関係の強弱の連続空間への操作.

==

—

5. 既存の DB, IR モデルに対する位置付け

5.2 Croft の IR モデルと指数型分布族のモデル

索引 X を持つ文書のある質問への関連度 $g(X)$ を

$$g(X) = \log \frac{P(X|\ Relevant)}{P(X|\ Nonrelevant)} = \sum_{i=1}^{n} x_i \log \frac{p_i(1-q_i)}{q_i(1-p_i)} + \sum_{i=1}^{n} \log \frac{1-p_i}{1-q_i},$$

とする (ベクトル IR : 索引 × 質問). 但し, それぞれ独立な, 質問に関連する文書 d が項 x_i を含む確率 p_i, 質問に関連しない文書 d が項 x_i を含まない確率 q_i とする.

⇒ $\boxed{\text{関連性 } p_i \text{ のみの IR を考えれば指数型}}$

$$\log p(x;\ \theta) = \sum_i \theta^i \log p_i(x) - \psi(\theta),$$

に帰着.

==

—

5. 既存の DB, IR モデルに対する位置付け

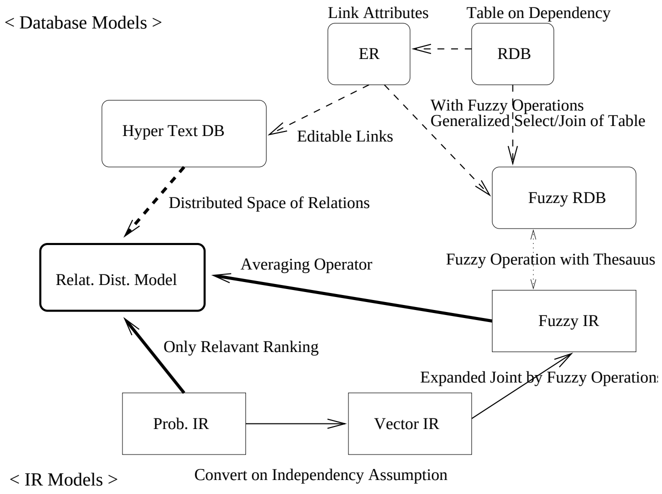

< Database Models >
Link Attributes
Table on Dependency
ER
RDB
Hyper Text DB
Editable Links
With Fuzzy Operations
Generalized Select/Join of Table
Fuzzy RDB
Distributed Space of Relations
Relat. Dist. Model
Averaging Operator
Fuzzy Operation with Thesauus
Fuzzy IR
Only Relavant Ranking
Expanded Joint by Fuzzy Operation
Prob. IR
Vector IR
< IR Models >
Convert on Independency Assumption

図 4: 関係分布モデルの関係操作性と平均化作用

6. 発見的な情報探索と視点の推定

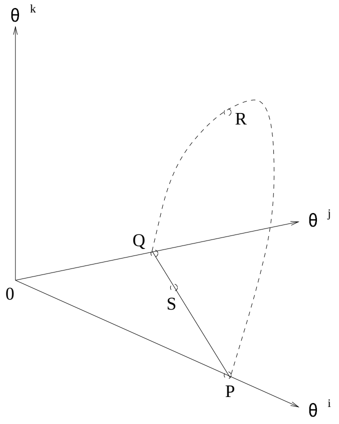


図 5: ±α-測地線上のデータ操作：探索範囲を示す 2 つの視点に対する補間的な操作（実線 α）と発見的な操作（点線 $-\alpha$）

⇒ 指数型の m-測地線は, 感覚量 $\log p_i(x)$ の期待値$\eta_i = \sum_{x\in\mathcal{X}} \log p_i(x)p(x;\ \theta)$ の補間に相当.

6. 発見的な情報探索と視点の推定

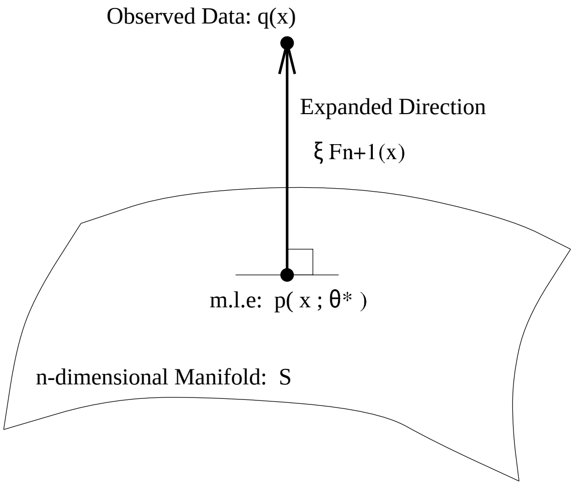

Observed Data: q(x)
Expanded Direction
ξ Fn+1(x)
m.l.e: p(x ; θ*)
n-dimensional Manifold: S

図 6: 最ゆう推定値θ^*と $p(x;\ \theta^*, \xi)$ による関数空間中のモデル拡張

$\Rightarrow$ 推定された視点θ^*と, 新たな関連度$p_{n+1}(x)$が得られる.

—

6. 発見的な情報探索と視点の推定

観測データ $q(x)$ から指数型分布族の多様体 $S=\{p(x;\ \theta)\}$ への最ゆう推定点θ^*を求めて, 多様体 Sの 1 次元拡張$\tilde{S}=\{p(x;\ \theta,\xi)\}$

$$p(x;\ \theta,\xi)=\exp\left[C(x)+\sum_{i=1}^{n}\theta^i F_i(x)+\xi F_{n+1}(x)\right],$$

を考え $(C(x) = p_0(x) \overset{\text{def}}{=} \frac{1}{m},\ F_i(x) \overset{\text{def}}{=} \log\frac{p_i(x)}{p_0(x)})$, 制約条件 $\sum_x p_{n+1}(x)=1$ のもとで, $f(y,\xi)$ を最小にする関数 $y_x \overset{\text{def}}{=} F_{n+1}(x)$ を求める.

$$min : f(y,\xi) \overset{\text{def}}{=} D(q(x)||p(x;\ \theta^*,\xi)) = const.-\xi\sum_{\mathcal{X}} q(x)y_x+\psi(\theta,\xi),$$

$$s.t : h(y) \overset{\text{def}}{=} m-\sum_{x=1}^{m}\exp(y_x)$$

==

—

6. 発見的な情報探索と視点の推定

拡張 Lagrange 関数 $L_r(y, \xi, \mu_k)$ を考えて乗数法で解く.

$$L_r = f(y, \xi) + \mu_k h(y) + \frac{1}{2} r_k \{h(y)\}^2$$

<ξを固定した場合>

任意の $y_x \in R$に対する関数 $f(y, \xi)$ の Hessian は, $\partial^2 f / \partial y_x^2 = \xi^2 p(x;\ \theta, \xi) > 0$, $\partial^2 f / \partial y_x \partial y_{x'} = 0\ (x' \neq x)$ より正定値となるので, $f(y, \xi)$ は下に凸な関数. また, $\partial^2 h / \partial y_x^2 = -\exp(y_x) < 0$, $\partial^2 h / \partial y_x \partial y_{x'} = 0$ より, $h(y)$ は上に凸な関数.

$\Rightarrow \boxed{\text{凸計画問題となって唯一の最適解 } y_x \text{ が存在}}$.

<ξも変数とした場合>

$\frac{\partial^2 f}{\partial \xi^2} = \sum_{\mathcal{X}} y_x^2 p(x;\ \theta, \xi) = E_{\theta,\xi}[y_x^2] > 0$ だが, $\frac{\partial^2 f}{\partial \xi \partial y_x} = \frac{\partial^2 f}{\partial y_x \partial \xi} = (1 + \xi y_x) p(x;\ \theta, \xi) - q(x)$ の符号は一定でないので, 関数 $f(y, \xi)$ の Hessian は正定値とは限らず, 多数の解が存在しうる.

===

7. まとめ

- 情報幾何学のα-分布族が, ファジィの平均演算子の作用を持つという, 両者のつながりを示した.
- α-分布族による DB モデルを提案し, 従来の DB モデルや IR モデルとの関係を整理した.
- 提案モデルにおける ($\pm\alpha$-測地線上の) 情報探索法と, 観測データに対する視点の推定法を紹介した.

⇒ 最適 / 妥当なαの値 (分布族 / 平均) や探索経路は, 応用分野ごとに評価 (人々の主観の評価) が必要だろう !!

==

人々の Recommend を活かした魅力的な情報空間

動的な関係構築の必要性: F. Halasz, Reflections on Note Cards: Seven Issues for the Next Generation of Hypermedia Systems, CACM'88

Personal/Collabo フィルタリング: D. Goldberg, et al., Using Collaborative Filtering to Weave an Information Tapestry, CACM'92.
S. Loeb, Architecting Personalized Delivery of Multimedia Information, CACM'92.

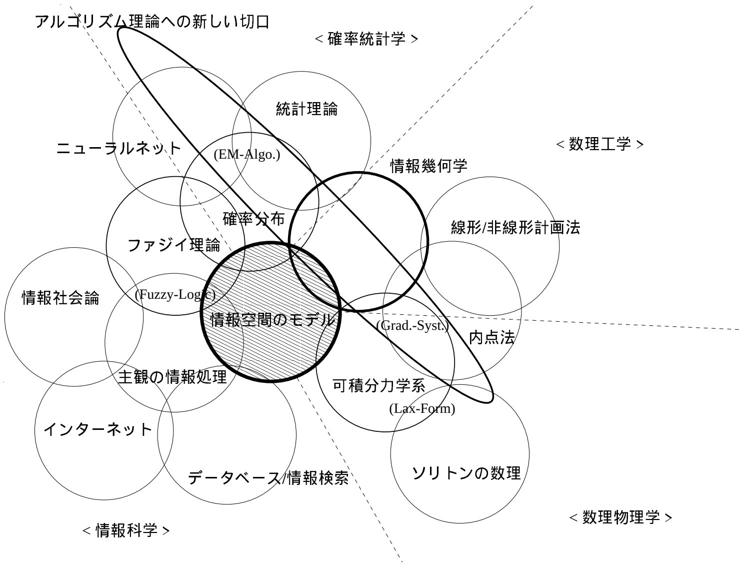
カスタム化の要望: C.M. Bowman, et al., Scalable Internet Resource Discovery: Research Problems and Approaches, CACM'94.

人的資源/推奨の活用: W. Hill, et al., Recommending and evaluating choices in a virtual community of use, CHI'95 Conf. Proc.'95.
U. Shardanand and P. Maes, Social Information Filtering: Algorithms for Automating "World of Mouth," CHI'95 Conf. Proc.'95.
M. Balabanovic and Y. Sholam, Content-Based, Collaborative Recommendation, CACM'97.

==

—

研究分野の広がり


アルゴリズム理論への新しい切口
< 確率統計学 >
統計理論
ニューラルネット
(EM-Algo.)
< 数理工学 >
情報幾何学
確率分布
線形/非線形計画法
ファジイ理論
情報社会論
(Fuzzy-Logic)
情報空間のモデル
(Grad.-Syst.)
内点法
主観の情報処理
可積分力学系
(Lax-Form)
インターネット
データベース/情報検索
ソリトンの数理
< 数理物理学 >
< 情報科学 >

平均値のファジィ積分による表現

クリスプ集合 $N_i = \{1, 2, \ldots, i\}$ に対するファジイ測度$\mu(N_k)$ を同定する割当関数

$$\mathcal{X}(\gamma, \frac{1}{n}\sum_{i=1}^{k} u_i^{\beta}) \stackrel{\text{def}}{=} \left[\frac{1}{n}\sum_{i=1}^{k} u_i^{\beta}\right]^{\gamma}, \ \ \gamma \stackrel{\text{def}}{=} \frac{1}{\beta} \in (0, \infty). \qquad (1)$$

を用いると, 平均値$\left[\frac{1}{n}\sum_{i=1}^{n} u_i(x)^{\beta}\right]^{\frac{1}{\beta}}$ は$0 \leq u_i < 1$ なる範囲で，

ファジイ積分: $\vee_{k=1}^{n+1}\left[h(k) \wedge \mu(N_k)\right]$

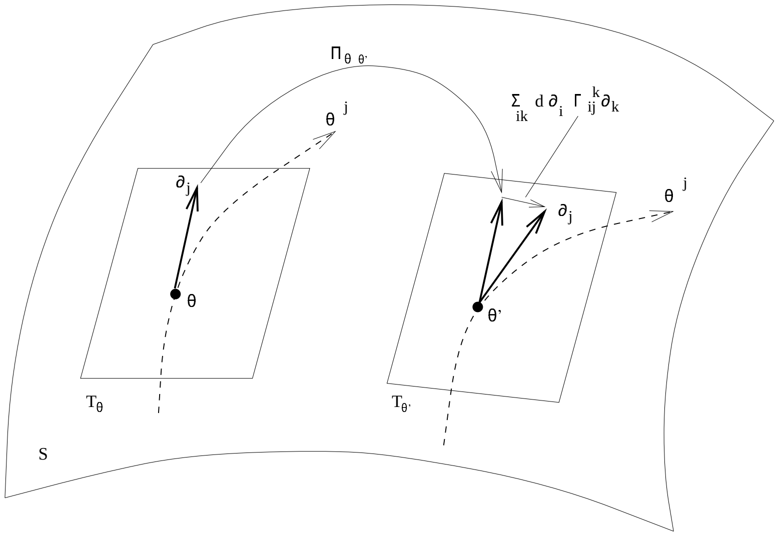
ショケ積分: $\sum_{k=1}^{n}(h_k - h_{k+1})\mu(N_k)$

によって表現できる（被積分関数： $h(k) \stackrel{\text{def}}{=} 1 \ (k = 1, 2, \ldots, n)$, $h(n+1) \stackrel{\text{def}}{=} 0$，ダミー変数： $u_{n+1}^{\beta} \stackrel{\text{def}}{=} n - \sum_{i=1}^{n} u_i^{\beta}$ ）.

$\mu(N_k)$は，$\beta > 1$のとき劣加法的(算術平均よりも最大値に近い)，
$0 < \beta < 1$のとき優加法的(算術平均よりも最小値に近い)．

==

双対平坦な統計的多様体 (S, g, ∇, ∇^*)


Π θ θ'
Σ ik d ∂ i Γ ij k ∂k
θ j
∂j
θ
T θ
θ j
∂j
θ'
T θ'
S

m 次元離散分布 $M = \{q(x)\}$ 中の n 次元部分多様体 $S = \{p(x;\ \theta)\}$

指数型分布族と混合型分布族

Exponential:

$$\begin{aligned} p(x;\ \theta) &= \exp\left[C(x)+\textstyle\sum_{i=1}^n \theta^i F_i(x) - \psi(\theta)\right], \\ &= p_0(x)\Pi_{i=1}^n \left[\frac{p_i(x)}{p_0(x)}\right]^{\theta^i} / \textstyle\sum_{x\in\mathcal{X}} \left\{p_0(x)\Pi_{i=1}^n \left[\frac{p_i(x)}{p_0(x)}\right]^{\theta^i}\right\}, \end{aligned}$$

Mixture: $p(x;\ \theta) = p_0(x) + \sum_{i=1}^n \theta^i(p_i(x) - p_0(x))$,

但し, $C(x) = \log p_0(x)$, $F_i(x) = \log \frac{p_i(x)}{p_0(x)}$とした.

$n = 2$ の例： m 次元の確率ベクトル $p, q \in M$を結ぶ 2 つの方法

$$\begin{aligned} r(x) = &\quad C_t \times p(x)^{1-t} q(x)^t, \\ s(x) = &\quad (1-t)p(x) + tq(x). \end{aligned}$$

時間変数 $0 < t < 1$, 規格化定数 $C_t = \{\sum_{x\in\mathcal{X}} p(x)^{1-t} q(x)^t\}^{-1}$.

==

—

指数型と混合型の $p(x;\ \theta)$ による要素 x のベスト 10

順位	指数型	混合型	$p_1(x)$	$p_7(x)$	$p_2(x)$	$p_5(x)$
1	*57*	68	68	84	41	26
2	26	26	87	83	24	22
3	24	24	88	33	16	15
4	16	16	*57*	92	*57*	13
5	2	22	44	70	51	*96*
6	68	15	38	*96*	99	8
7	*96*	13	37	90	98	65
8	41	41	*96*	89	97	63
9	34	84	94	66	56	59
10	29	83	86	47	42	34

p_1, p_7 と p_2, p_5 の両者で上位の $x = \mathit{57}, \mathit{96}$ が高位にある指数型の AND 的作用と, 両者の上位を混在させた混合型の OR 的作用.

インタラクションの拡張

	物理世界	記号世界	情報世界
対話の相手	物理的実在	抽象的実在（コード）	抽象的認知物
対話の結果	外界への動作	計算, 文書編集	概念理解, 状況把握
延長機能	身体	記号処理	直観的思考？
操作感覚	外的	内的	深的
道具の例	斧, 工作機械, 車	文書エディタ	次世代 DB？

太古の人類から現代に至るまでの科学技術の発展によって,
物とのインタラクション → 記号とのインタラクション→ 情報
や知識とのインタラクションへと拡張しつつある.

==

魅力的な情報は？

客観的なもの（ロゴス）	主観的なもの（パトス）
歴史的な事実	評価が未定なトピックス
辞書	小説, 記事
法律, 規則	経験などに基づく諺
観測データの羅列	さまざまな意見の集約
商業メディアを介した公式情報	口コミ情報
予め固定化された体系, スキーマ	視点に応じた関係構造
演繹や論理による関係	洞察や直観による関係
受動的な意味伝達	対話的な意味解釈

⇒ （他人の）さまざまな視点から吟味された情報をもとに, 自らの目的や意図に合った活用を見い出せること, すなわち, 現実世界に活かせる（生きた）意味づけが要求される.

===