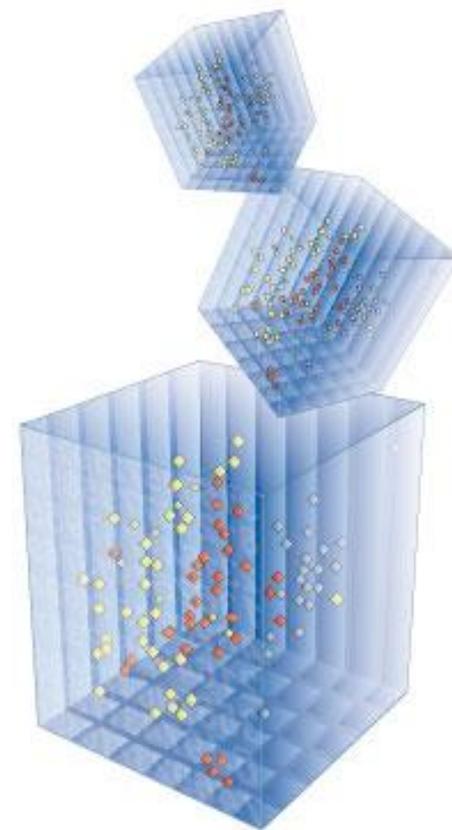


ビッグデータ時代の One to oneマーケティングの実際

2014年2月4日（火） 10:30 - 11:00
於：タワーホール船堀

お問い合わせは

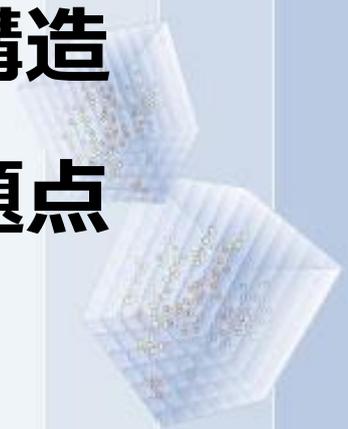
y_yamakawa@mcx.co.jp



代表取締役会長 山川義介

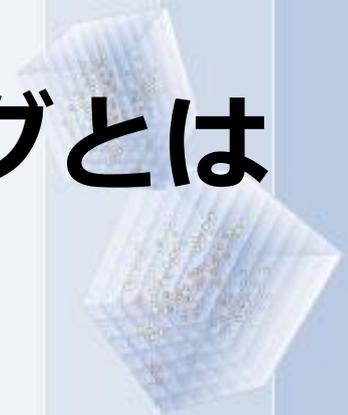
アルベルト
株式会社ALBERT

〒151-0053 東京都渋谷区代々木2-22-17
TEL:03-5333-3703 (代表) FAX:03-5333-3723
www.albert2005.co.jp/

- 
1. One to oneマーケティングとは
 2. ビッグデータ（購買データ）の構造
 3. 従来のクラスター分析手法と問題点
 4. ALBERT距離の提案
 5. マーケティング施策への応用
- 



1. One to oneマーケティングとは



1



to 1 顧客



実務的には！

to



1.人口統計的データ

性別、年齢、居住地、職業、年収・・・（これによるセグメントは崩壊しつつある）

2.心理的データ

価値観、ライフスタイル・・・（心理的データからの購買予測は極めて困難）

3. データ

閲覧履歴、カート投入履歴、購買履歴・・・（現在もっとも多く使われている）

4.コミュニケーションデータ

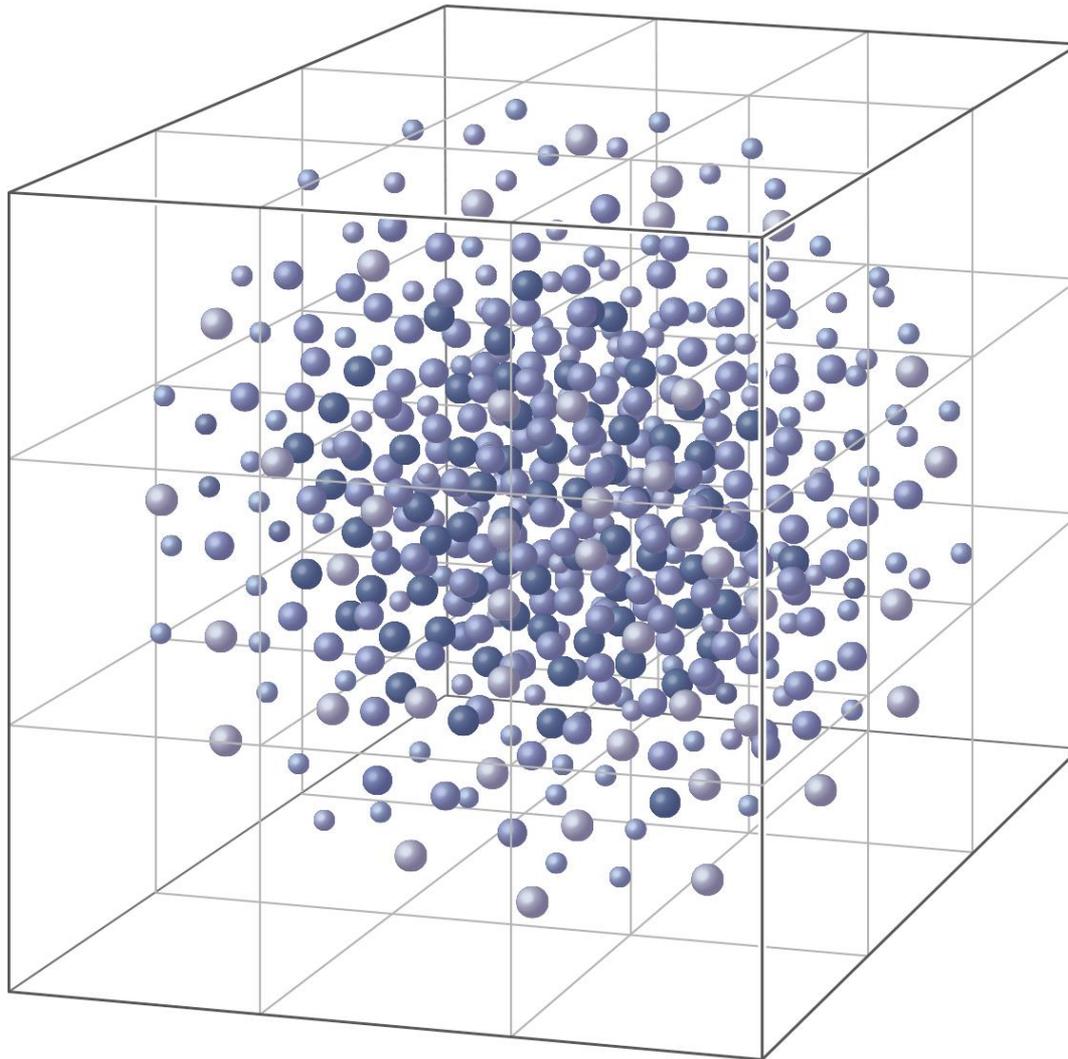
情報収集、質問、クレーム、要望、ブログ・・・（まだ十分な活用がされていない）

2.ビッグデータ（購買データ）の構造

アンケートデータ

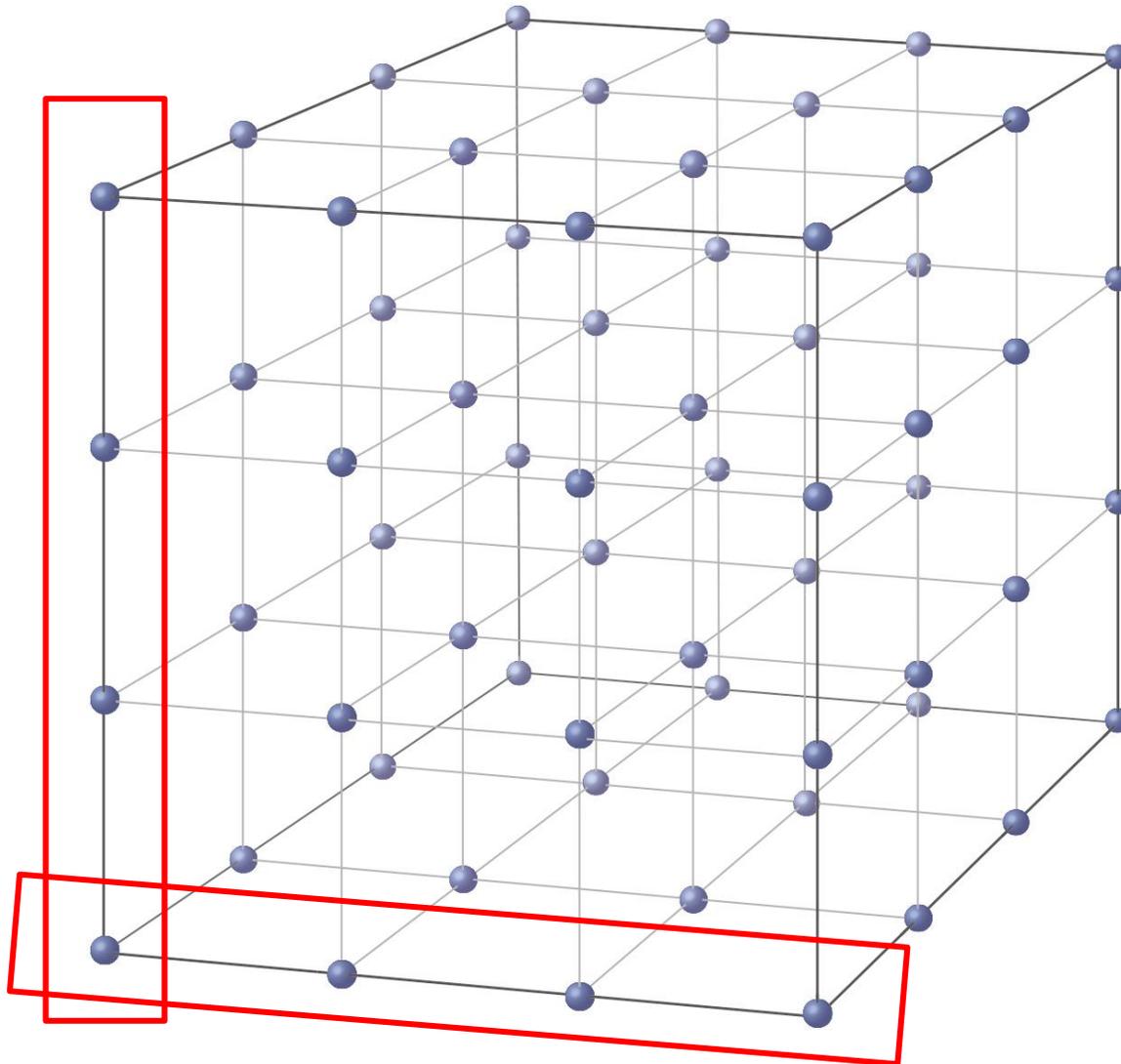
	回 答 1	回 答 2	回 答 3	回 答 4	回 答 5	回 答 6	回 答 7	回 答 8	回 答 9	回 答 10	回 答 11	回 答 12	回 答 13	回 答 14	回 答 15	回 答 16	回 答 17	回 答 18	回 答 19	回 答 20	回 答 21	回 答 22	回 答 23	回 答 24	回 答 25	回 答 26	回 答 27	回 答 28	回 答 29	回 答 30	回 答 31	回 答 32	回 答 33	回 答 34	回 答 35	回 答 36	回 答 37	回 答 38	回 答 39	
sample1	5	4	2	2	5	3	4	1	2	2	2	4	2	1	4	3	5	5	4	4	5	2	3	3	5	4	4	5	5	3	2	5	3	1	5	3	3	5	1	
sample2	1	3	2	5	2	2	5	5	2	2	2	1	5	3	1	2	3	2	2	5	1	3	1	3	5	4	2	4	4	2	3	1	2	3	2	5	2	3	4	4
sample3	4	2	4	2	3	3	4	4	3	4	3	3	3	1	2	5	3	3	4	1	2	5	3	3	2	4	1	2	3	2	3	1	3	2	4	3	1	3	5	5
sample4	4	2	3	4	2	3	2	2	2	5	4	2	4	2	3	2	2	3	1	3	4	3	1	4	1	3	5	4	2	3	4	1	1	4	3	1	5	3	4	4
sample5	1	3	1	3	4	4	1	3	3	3	4	1	1	2	2	2	2	5	4	3	5	4	4	2	4	3	2	5	3	4	3	3	1	2	2	3	4	2	4	4
sample6	4	2	1	2	5	5	2	2	3	4	3	3	4	2	4	5	3	2	2	4	2	4	4	2	5	4	5	4	4	3	5	1	3	3	1	3	3	4	3	3
sample7	5	2	5	2	3	5	4	2	3	3	4	4	4	4	3	4	4	4	3	4	3	4	2	1	4	2	1	4	5	3	2	4	2	5	5	1	1	2	2	2
sample8	2	2	4	5	5	3	2	4	2	4	4	4	4	3	3	5	5	3	3	1	4	1	4	3	1	4	3	2	1	4	5	3	1	3	2	4	3	4	5	4
sample9	3	2	4	4	2	4	3	4	1	3	4	3	2	2	3	2	2	3	1	3	2	3	2	3	2	3	4	4	2	3	4	1	2	2	3	1	1	2	3	3
sample10	3	5	2	4	1	4	4	3	2	5	4	5	2	4	4	3	4	2	1	5	5	4	2	4	2	4	4	2	1	3	2	5	3	4	1	4	3	2	2	2
sample11	4	4	2	5	1	2	1	4	2	5	3	3	4	3	1	4	2	5	3	2	3	2	3	1	4	5	4	3	4	4	3	5	3	2	4	2	2	5	4	4
sample12	3	5	4	3	2	2	4	2	5	4	4	2	3	2	2	3	5	2	1	2	2	1	1	3	3	2	4	1	2	2	5	4	5	4	4	3	2	4	4	4
sample13	2	5	3	5	2	2	4	2	5	2	3	3	1	3	2	3	4	3	3	2	2	5	3	4	4	1	1	5	5	3	3	2	3	2	4	2	4	4	4	5
sample14	2	3	4	1	1	5	2	4	5	4	5	4	2	5	4	4	3	2	4	2	5	5	4	4	4	3	5	4	5	3	4	5	1	3	4	1	2	4	3	3
sample15	3	3	4	2	5	2	2	5	2	4	3	1	5	5	1	3	5	2	3	5	3	3	4	1	4	4	2	4	4	3	5	1	3	2	1	5	2	3	2	2
sample16	2	4	2	3	3	4	4	4	3	2	3	3	3	4	4	2	4	1	2	1	2	1	1	3	3	3	2	2	2	4	4	3	4	4	5	3	3	1	5	5
sample17	5	4	5	3	2	4	2	5	4	3	4	3	1	2	2	2	5	3	4	3	1	4	3	1	2	4	2	1	2	1	4	3	5	4	1	5	3	3	3	3
sample18	3	4	4	5	2	2	3	3	3	1	4	1	3	1	5	5	5	1	2	2	3	2	5	5	4	5	4	4	4	2	4	3	4	5	1	3	2	4	4	4
sample19	4	5	3	4	5	5	2	4	4	5	2	2	2	3	5	3	5	3	4	1	1	2	3	3	3	4	2	3	3	5	2	3	4	3	5	4	1	5	2	2
sample20	5	2	3	4	4	3	2	2	2	3	2	2	4	5	4	2	2	3	4	4	4	2	5	3	2	3	4	2	3	4	2	1	4	4	1	3	2	1	5	5
sample21	5	1	4	3	1	1	3	3	4	4	1	5	2	2	2	4	3	5	2	3	4	2	1	1	4	1	4	2	2	5	3	4	4	5	5	2	4	4	3	3
sample22	2	4	1	4	4	4	3	3	1	4	2	4	3	5	5	5	4	3	5	5	4	3	1	1	2	1	1	4	2	5	4	3	2	1	3	3	3	2	1	
sample23	3	1	2	2	2	1	4	4	4	4	4	3	3	5	5	2	3	2	3	2	2	5	4	2	4	3	1	2	5	3	2	2	2	4	2	4	1	5	5	5

アンケートデータのイメージ



購買データのイメージ

購入データがある！



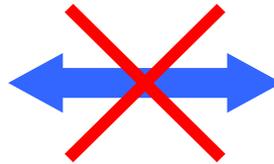
1.過去の購買履歴データの有効活用

2.データのsparse性対策



パンパース コットンケア ウルトラ
ジャンボ S 104枚 パンパース

ASIN: B0015XN55S

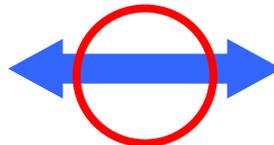


アサヒ スーパードライ
350ml×24缶

ASIN: B001TZAWD0

SKU単位の相関より**カテゴリ単位**の相関のほうがはるかに
パワフルで精緻な購買予測が可能になる。←非常に重要な視点

おむつ



ビール

「カテゴリ」に加え、同じ上位概念である「テイスト」「ブランド」の分析により顧客の理解が深まる。

Category

大分類、小分類



Taste

色、模様、サイズ



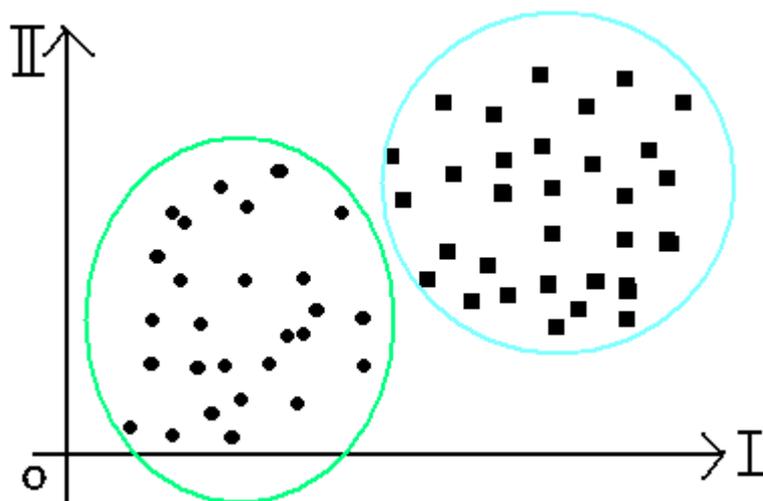
Brand

ブランド、キャラクター



3. 従来のクラスター分析手法と問題点

とにかく似ているものを集めてクラスターを作る。



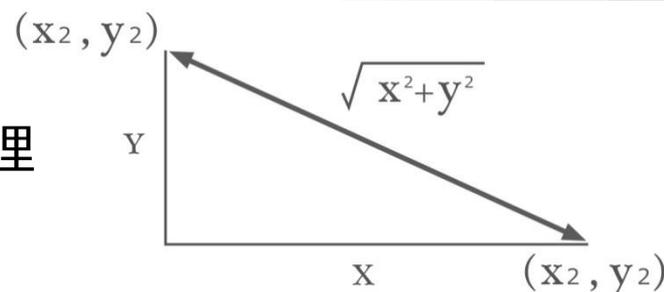
代表的手法は k-means法

今回の研究は、非階層型クラスター分析の中で、単純でわかりやすいため現在幅広く使われている、k-means法非階層型クラスター分析を用いる。

- ユークリッド距離

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2}$$

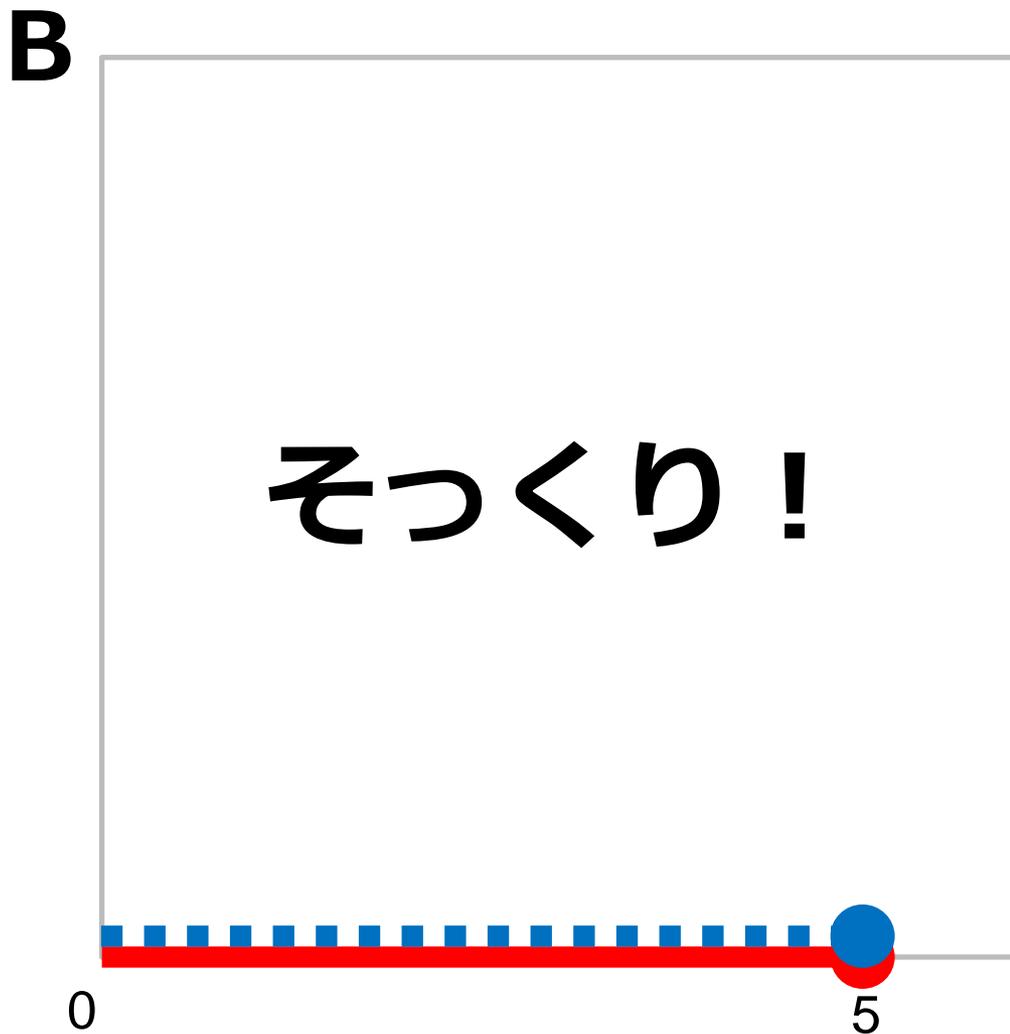
2次元の場合はピタゴラスの定理



- コサイン類似度 $= \frac{\vec{q} \cdot \vec{d}}{|\vec{q}| |\vec{d}|} = \frac{\vec{q}}{|\vec{q}|} \cdot \frac{\vec{d}}{|\vec{d}|} = \frac{\sum_{i=1}^{|V|} q_i d_i}{\sqrt{\sum_{i=1}^{|V|} q_i^2} \cdot \sqrt{\sum_{i=1}^{|V|} d_i^2}}$

距離への変換のため、ここでは $(1 - \text{cos類似度})$ をcos距離と定義しています。

両方とも5個買っている人同士



ユークリッド距離

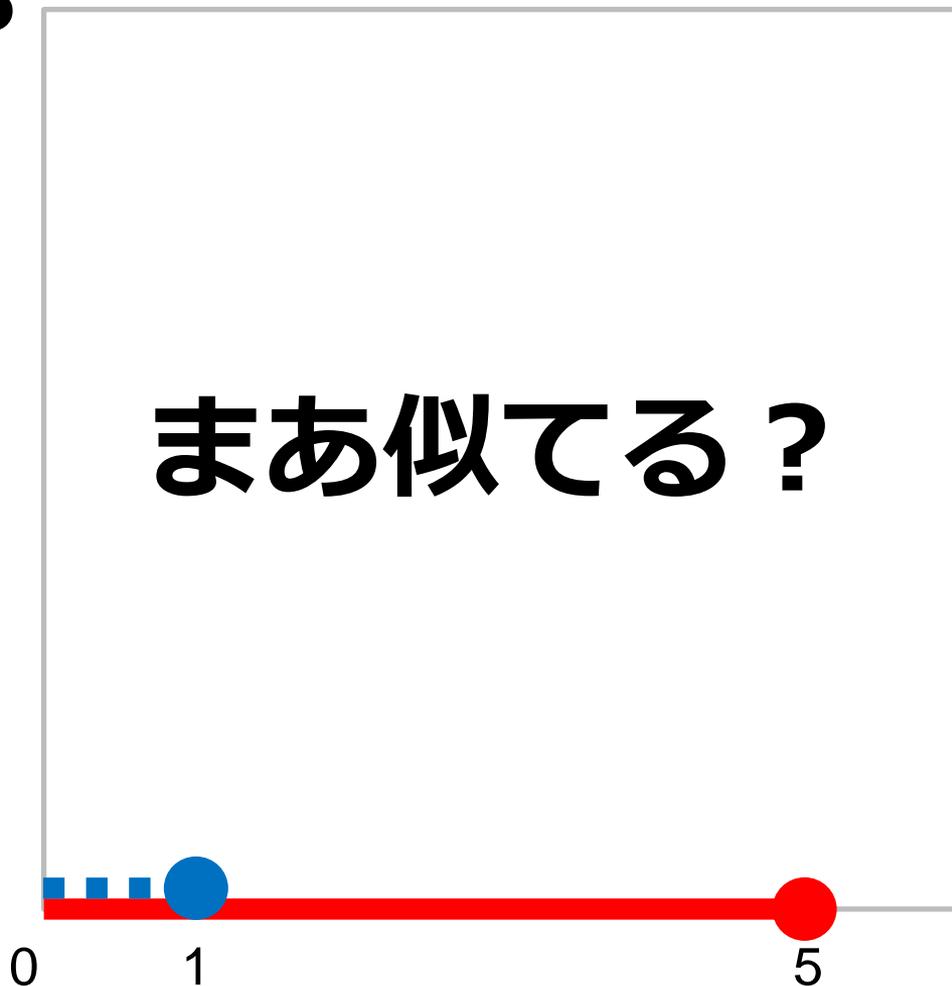
0

(1 - COS) 距離

0

Aを1個しか買っていない人と5個買っている人

B



ユークリッド距離

4.00

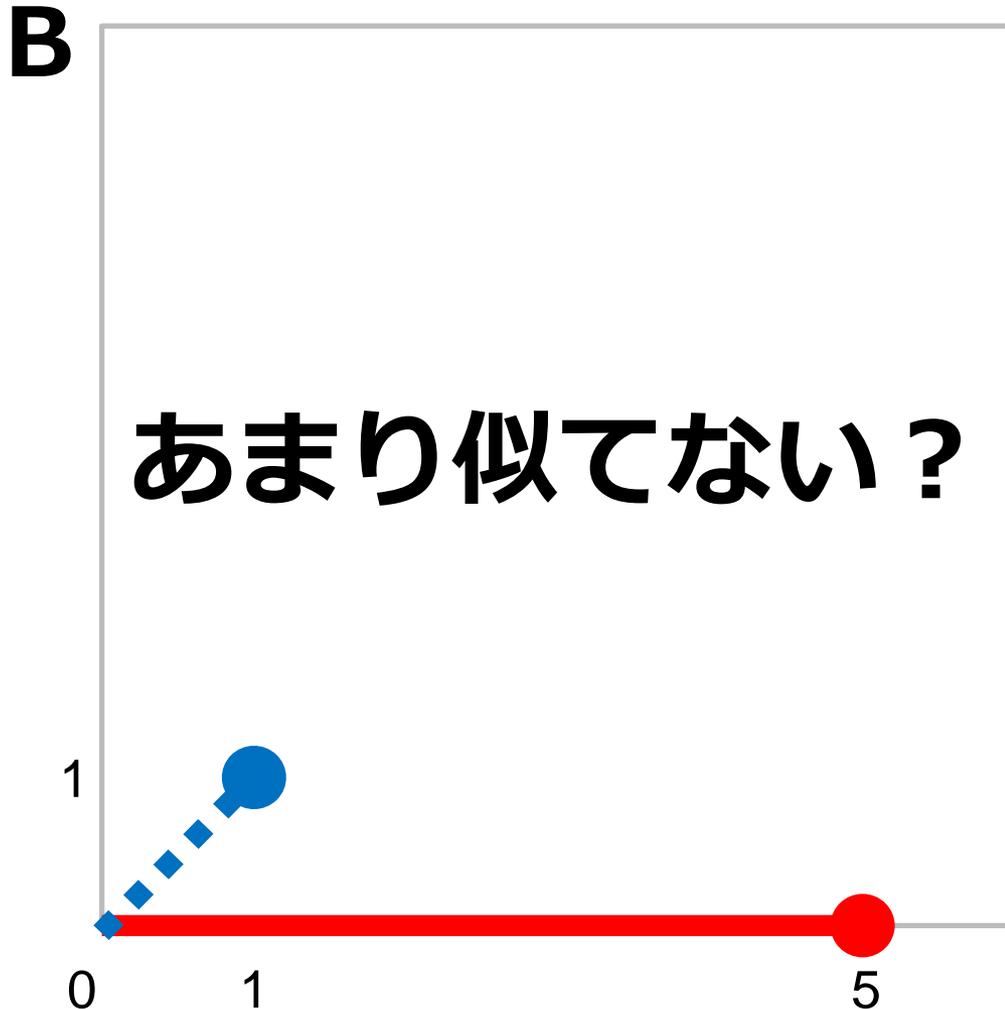
(1 - COS) 距離

0

A

従来の距離計算への疑問

AもBも1個ずつ買っている人とAだけ5個買っている人



ユークリッド距離

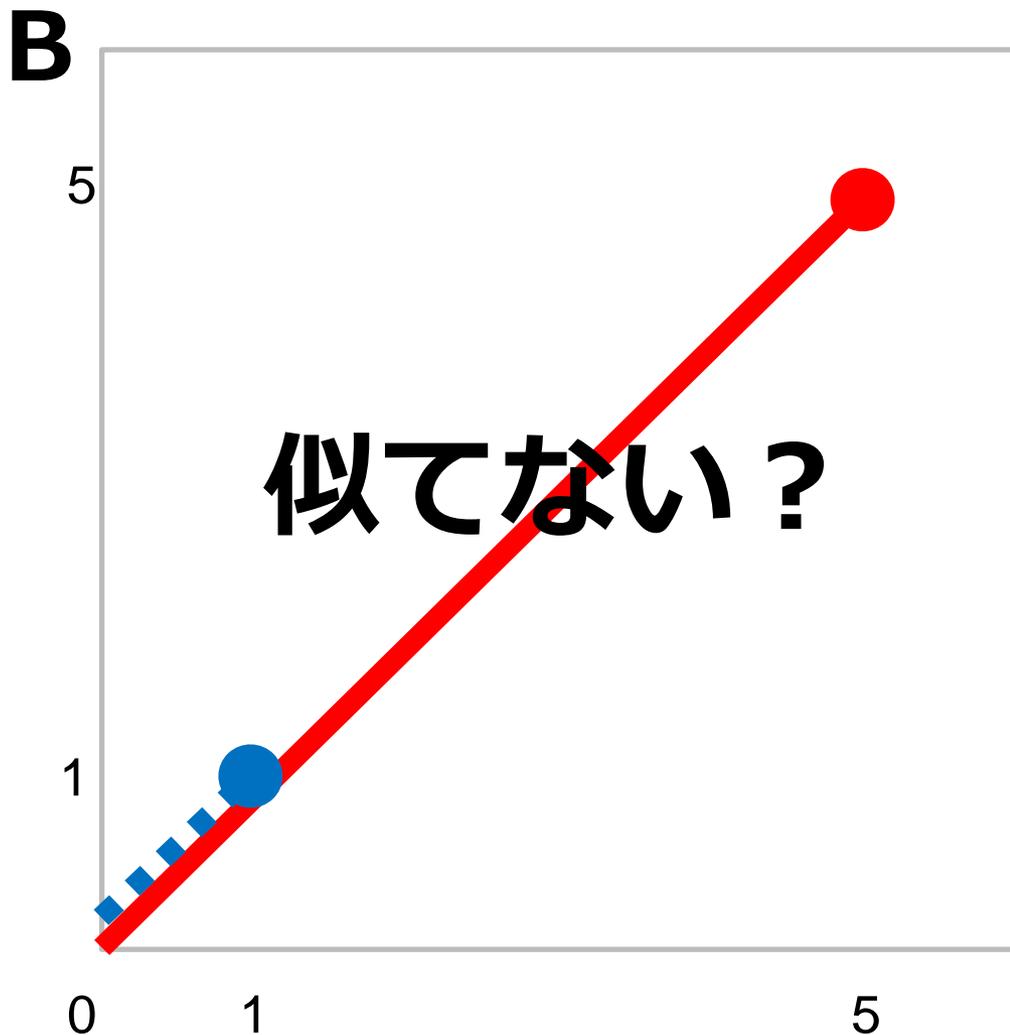
4.12

(1 - COS) 距離

0.293

従来の距離計算への疑問

AもBも1個ずつ買っている人と両方5個買っている人



ユークリッド距離

5.66

(1 - COS) 距離

0

1と0を入れ替えただけの購買パターン

caseP	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10		
s1	0	0	0	1	1	1	1	1	1	1	(1-cos) 距離	0.118
s2	1	0	1	1	1	1	1	1	1	1	ユークリッド距離	1.414

caseQ	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10		
s3	1	1	1	0	0	0	0	0	0	0	(1-cos) 距離	0.423
s4	0	1	0	0	0	0	0	0	0	0	ユークリッド距離	1.414

ユークリッド距離では同じ！

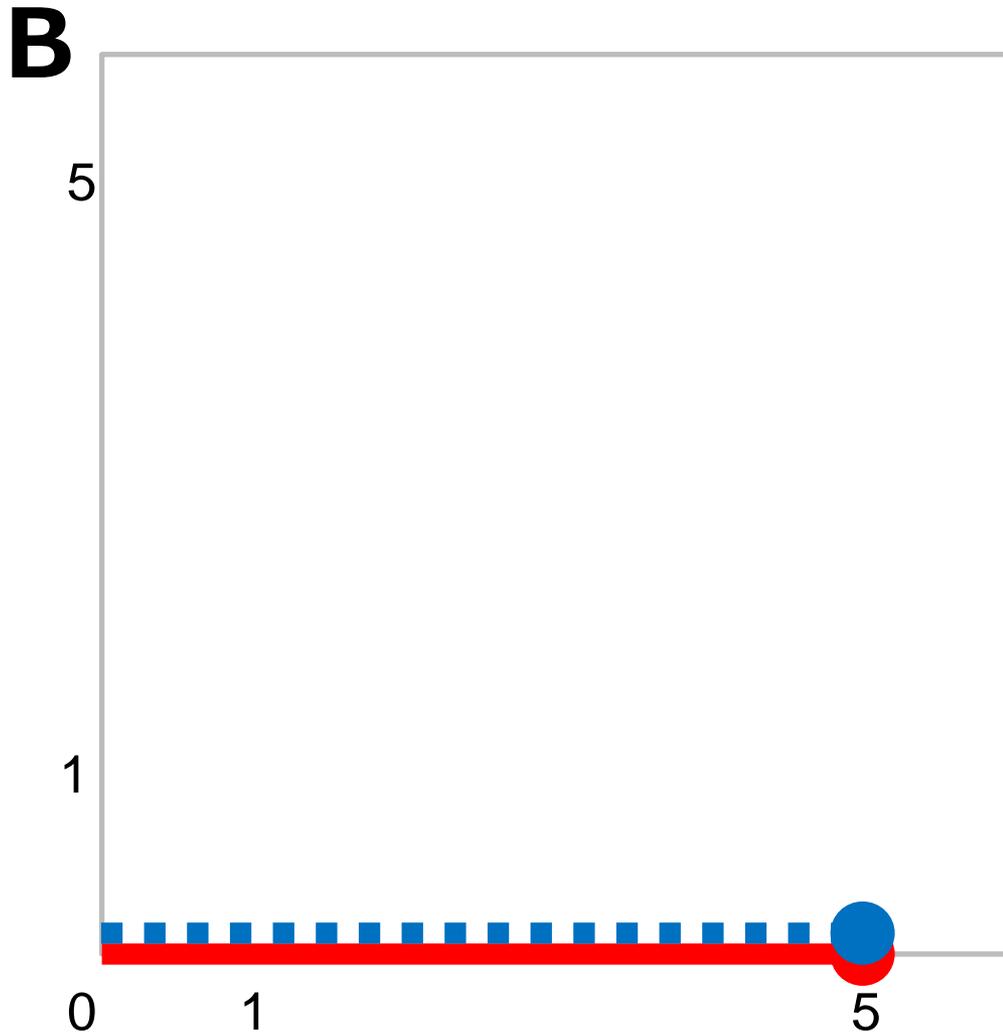
4. ALBERT距離の提案

(特許出願準備中につき詳細未公表)

1. 同じカテゴリを同数購入していても、
購入数によって類似度に重み付けをする。
2. 同じカテゴリを異なる数購入している場合、
その差によって類似度に重み付けをする。
3. 一方しか購入していないカテゴリに関して、
非類似度の概念を導入する。
4. 双方が購入していない場合は評価しない。
5. 類似度、非類似度はその寄与度をチューニング
できるようにパラメータ化する。

※同一点の距離が0にならるので、距離の公理という意味では距離ではありません。

両方とも5個買っている人同士



ユークリッド距離

0

(1 - COS) 距離

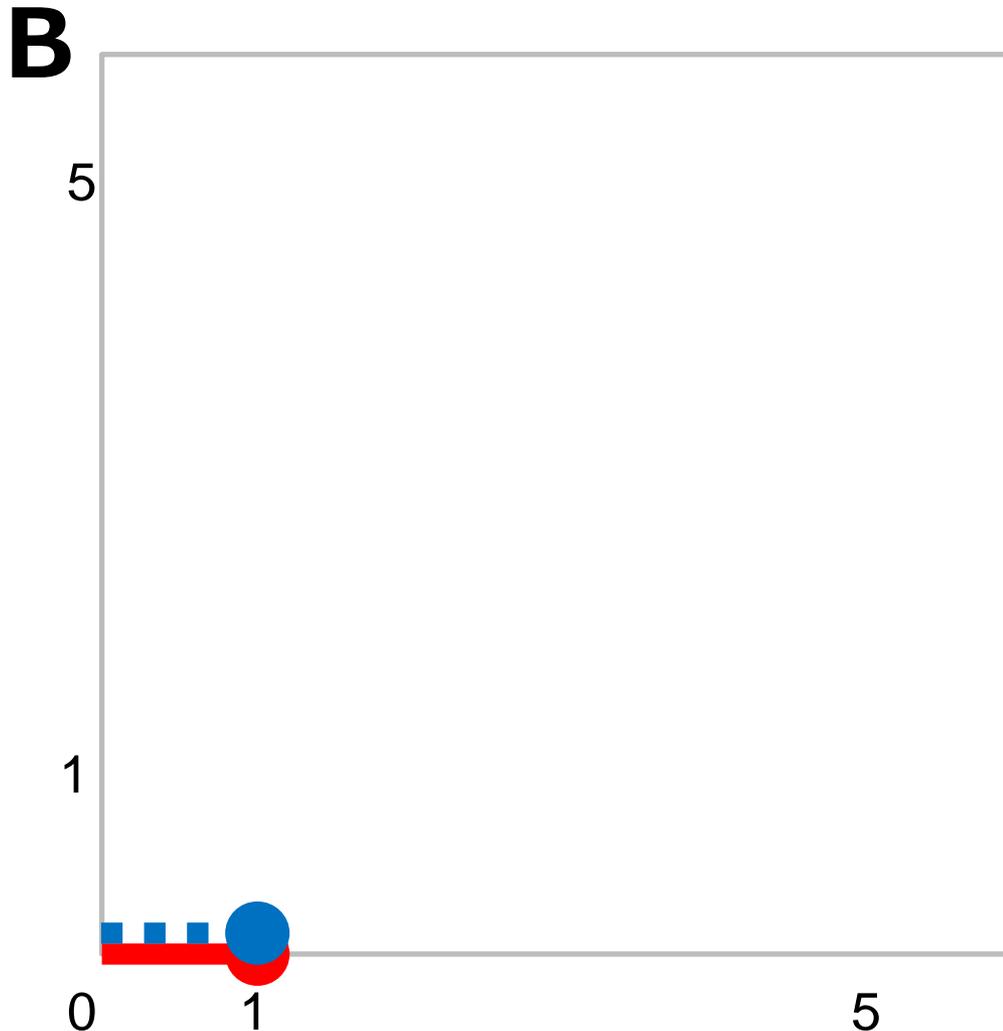
0

ALBERT距離

0.607

A

両方とも1個買っている人同士



ユークリッド距離

0

(1 - COS) 距離

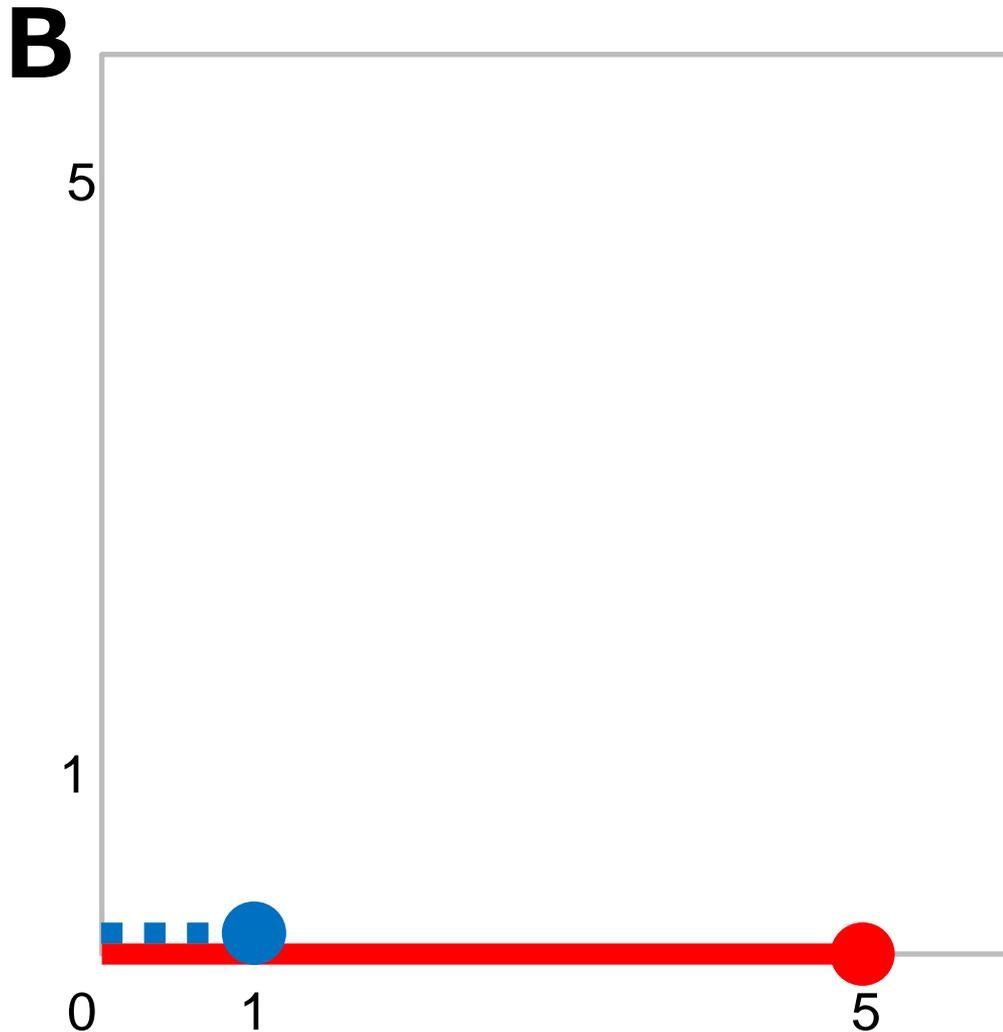
0

ALBERT距離

0.905

A

Aを1個しか買っていない人と5個買っている人



ユークリッド距離

4.00

(1 - COS) 距離

0

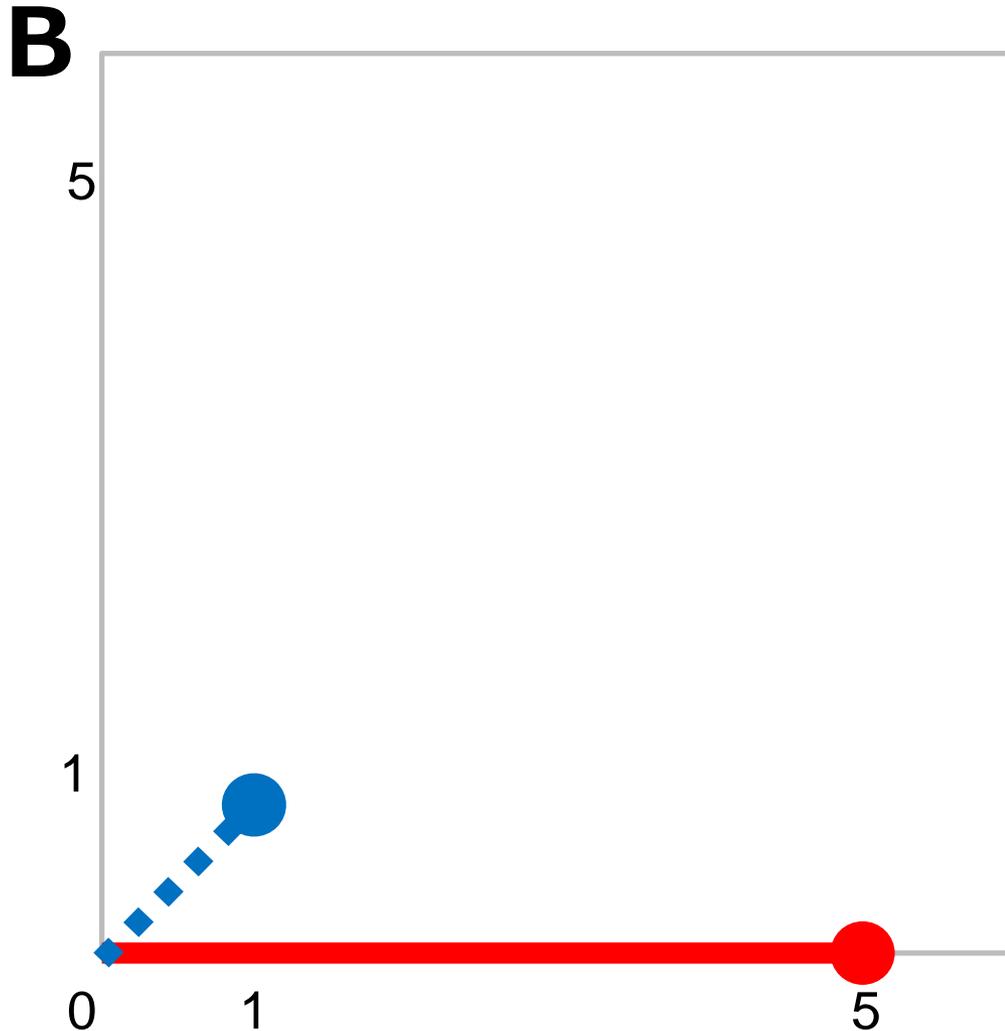
ALBERT距離

0.905

A

ALBERT距離の導入

AもBも1個ずつ買っている人とAだけ5個買っている人



ユークリッド距離

4.12

(1 - COS) 距離

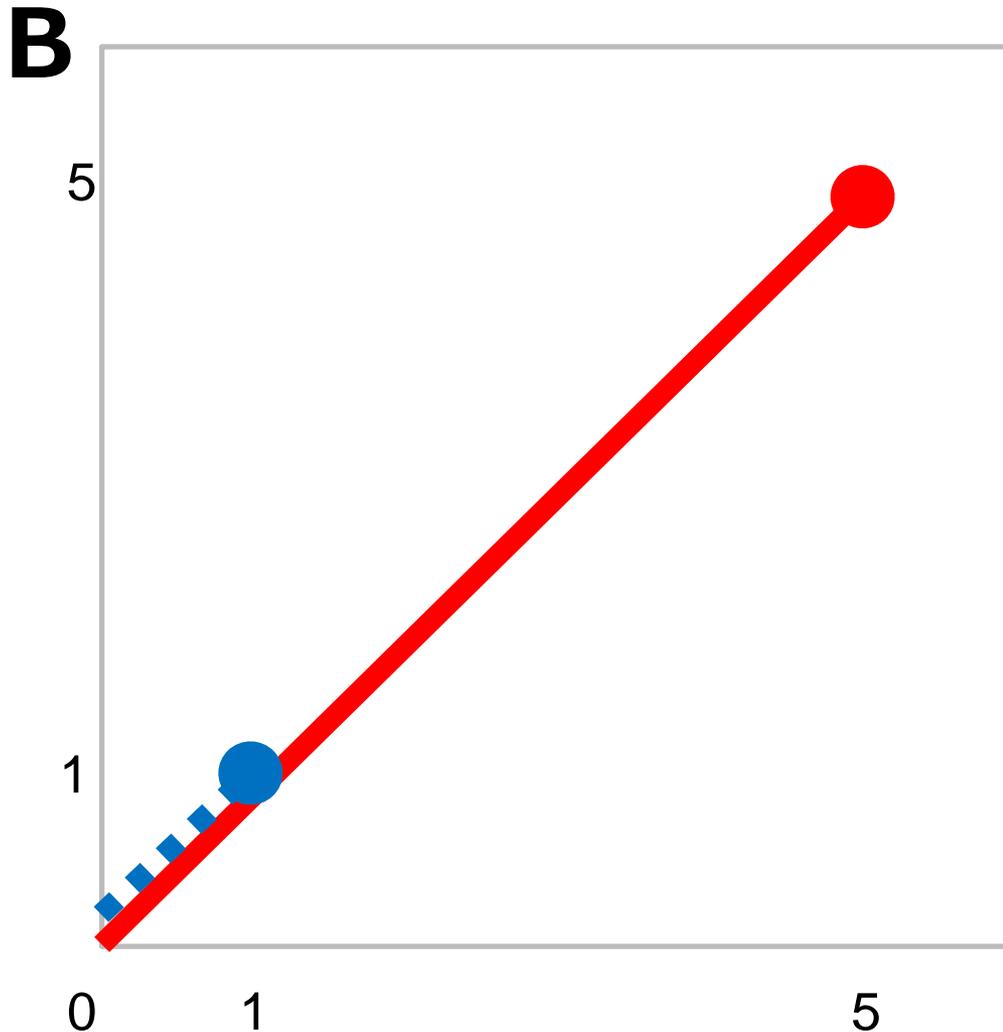
0.293

ALBERT距離

0.919

ALBERT距離の導入

AもBも1個ずつ買っている人と両方5個買っている人



ユークリッド距離

5.66

(1 - COS) 距離

0

ALBERT距離

0.819

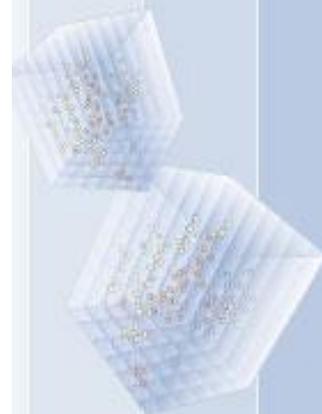
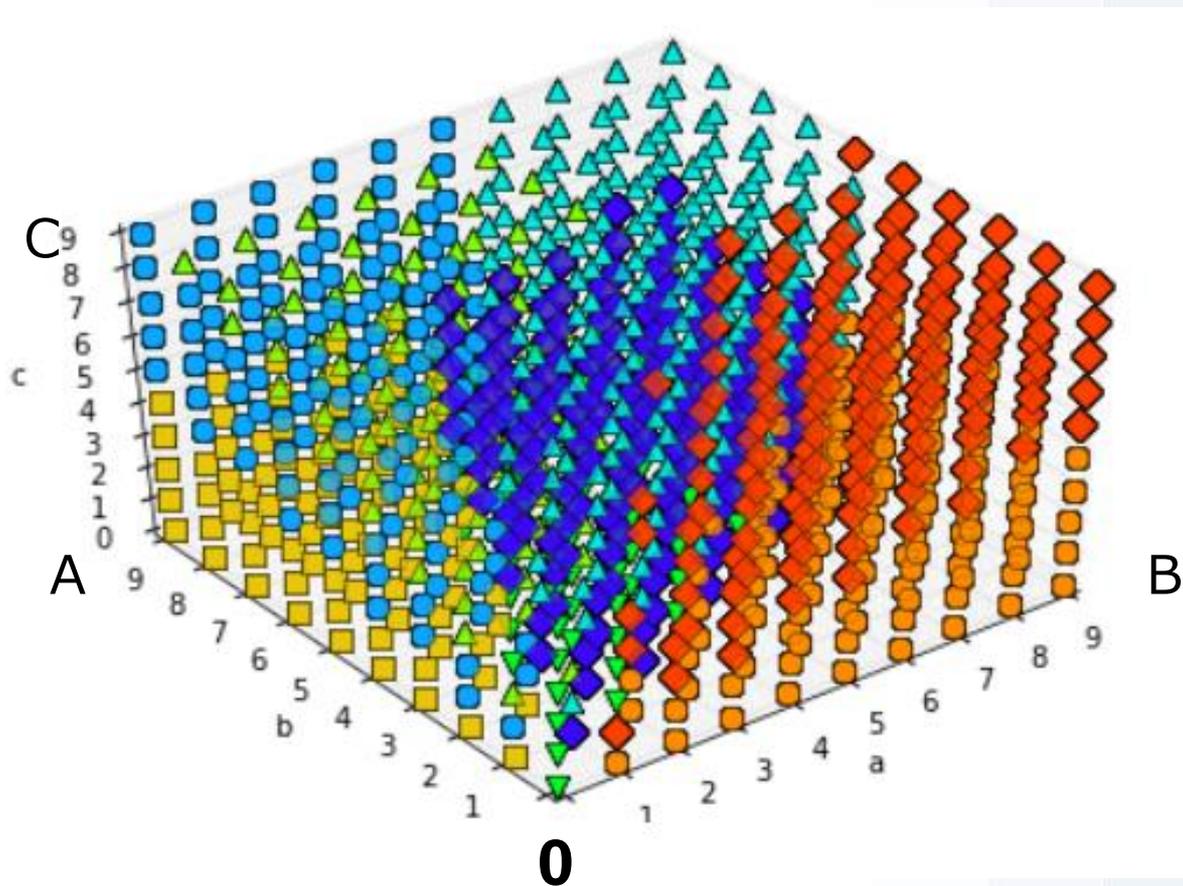
A

1と0を入れ替えただけの購買パターン

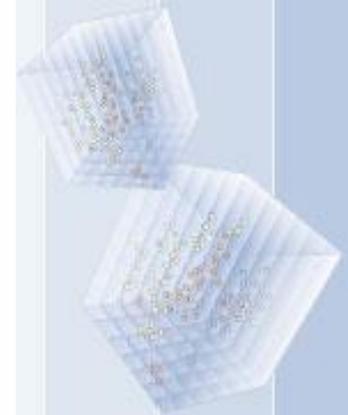
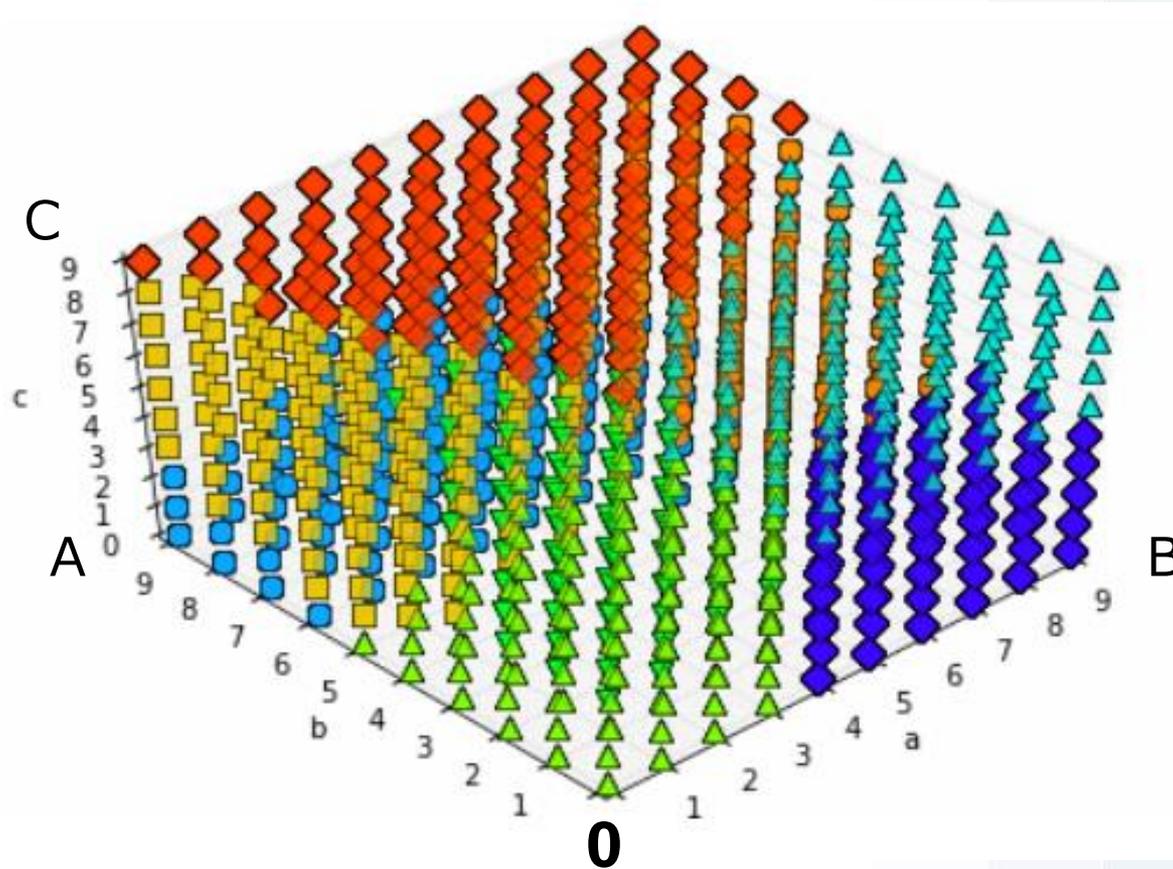
caseP	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10		
s1	0	0	0	1	1	1	1	1	1	1	(1-cos) 距離	0.118
s2	1	0	1	1	1	1	1	1	1	1	ユークリッド距離	1.414
類似度	-0.2	0	-0.2	0.5	0.5	0.5	0.5	0.5	0.5	0.5	ALBERT距離	0.726

caseQ	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10		
s3	1	1	1	0	0	0	0	0	0	0	(1-cos) 距離	0.423
s4	0	1	0	0	0	0	0	0	0	0	ユークリッド距離	1.414
類似度	-0.2	0.5	-0.2	0	0	0	0	0	0	0	ALBERT距離	0.980

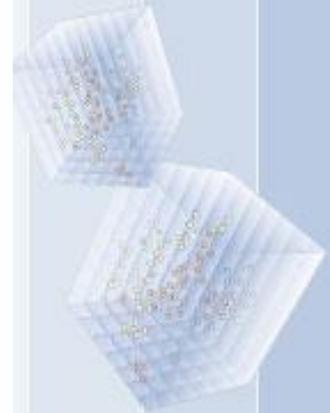
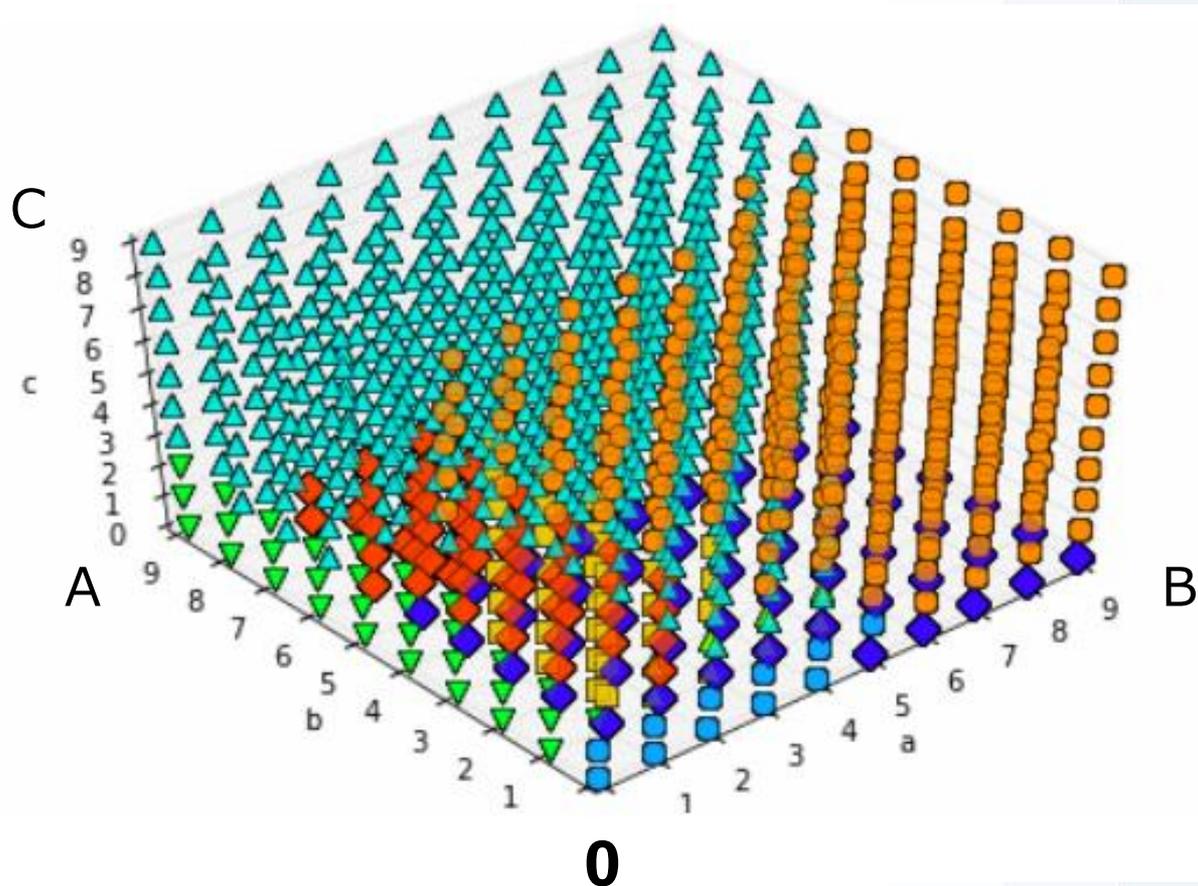
空間が放射線上に分割される傾向



空間が均等に分割される傾向

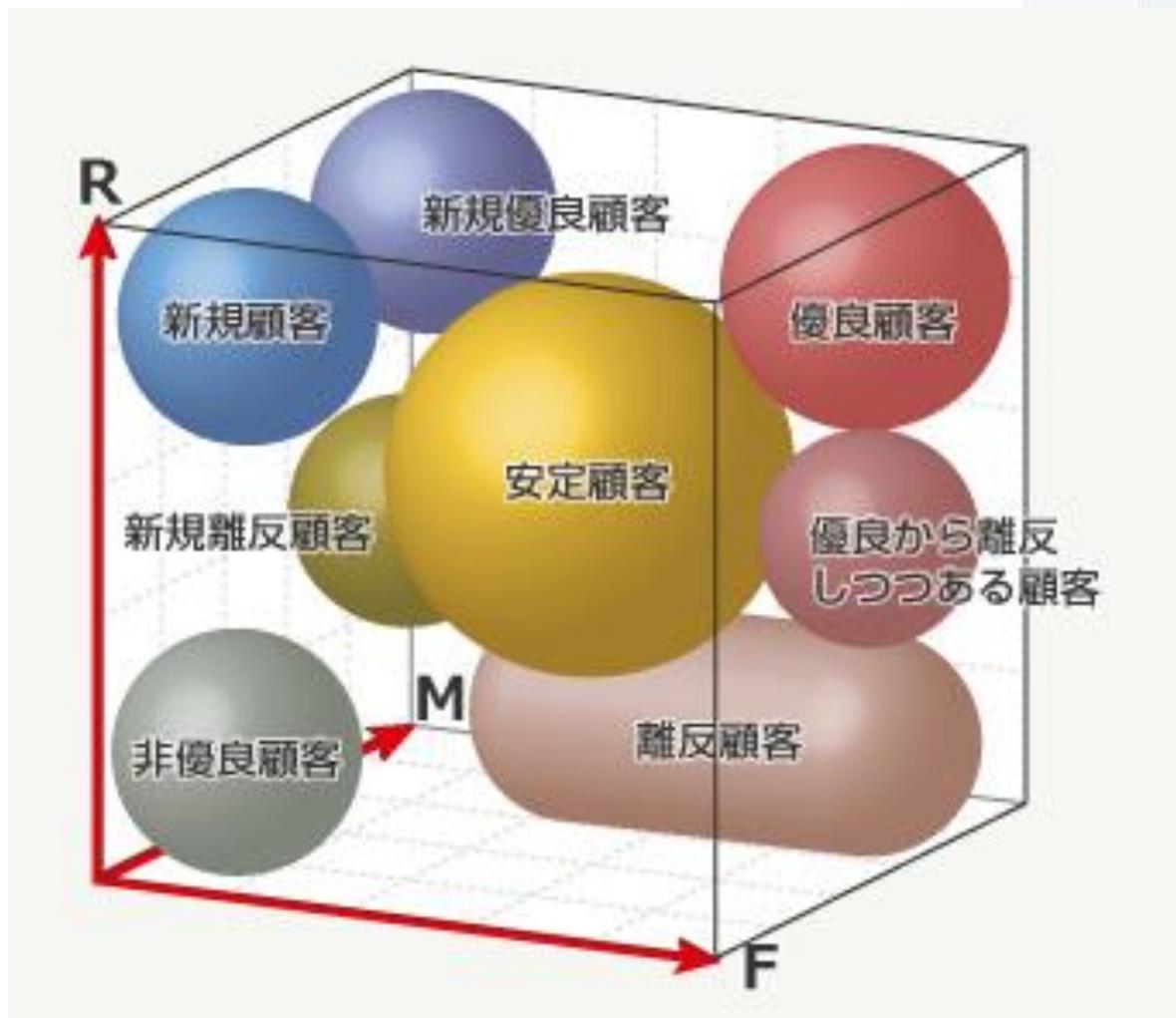


軸や平面付近とそれ以外に分割される傾向

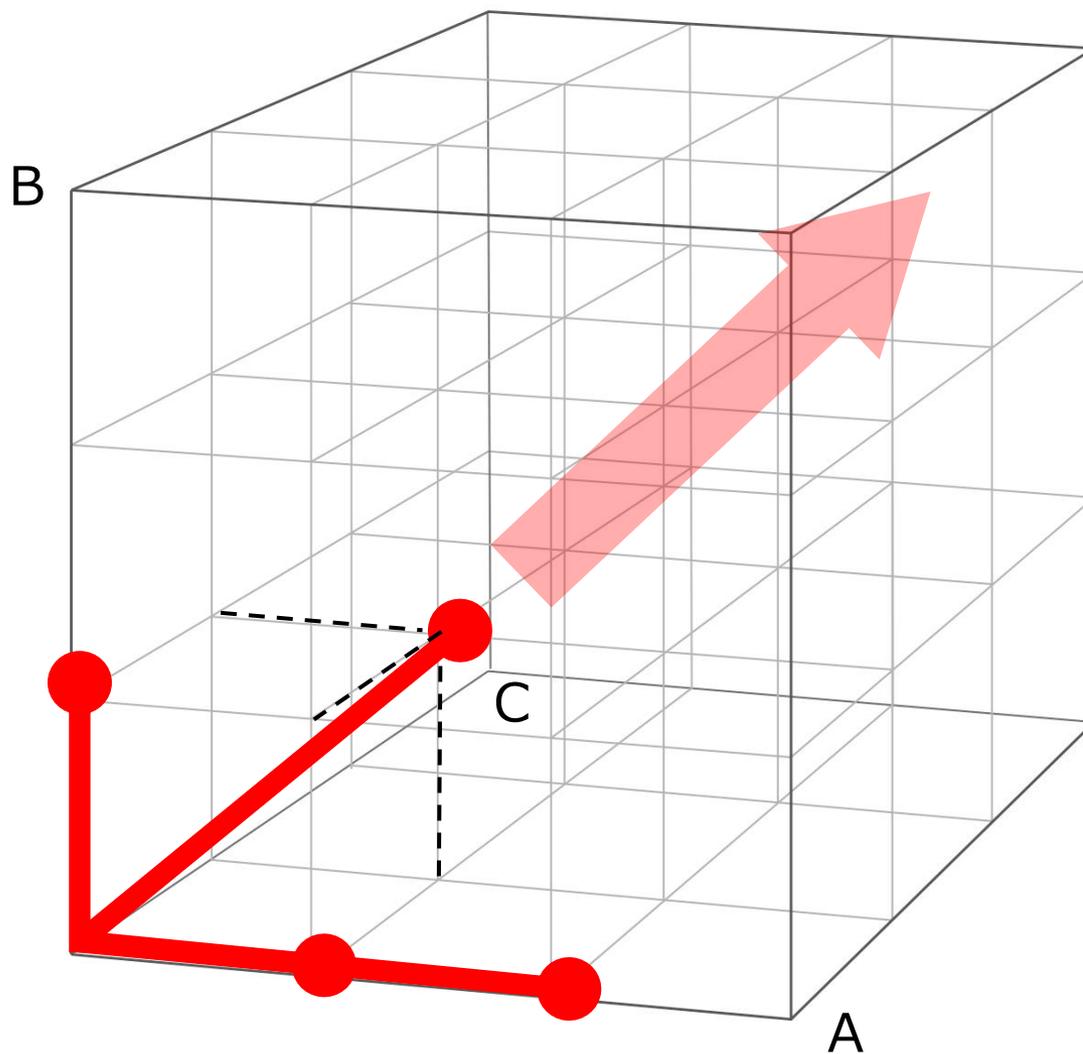


5.マーケティング施策への応用

RFM分析におけるロイヤルカスタマー



購入カテゴリを考慮したロイヤルカスタマー化



朝野熙彦 編著



ビッグデータの 使い方・活かし方

マーケティングにおける活用事例

- ALBERT
- インデックス・アイ
- インサイト・ファクトリー
- ビデオリサーチ
- リクルートライフスタイル
- R&D
- KSP-SP

データサイエンス
のトップランナーが紹介する最先端の
情報抽出・情報解析 **10選!**

ビジネスに役立つ情報を発見!

東京図書

山川義介 (株)ALBERT

第1章

ビッグデータ時代のレコメンデーション

POINT

- ビッグデータの情報活用で、最も顕著な成果を上げているのが
通販サイトのレコメンデーション機能である
- 多種多様なレコメンデーションを分類するフレームとして、
ACKマトリックスを提案した
- 単純な共起性ではなく情報量を加味したレコメンデーションのロジックが必要である
- 解析技術的には変数が膨大な割にはデータが少ないというスパースなデータの
解析が問題になるが、本章ではその解決策を示した
- 千趣会のベルメゾンネットでは顧客のクラスター分析を行うことで、
レコメンデーションのパーソナライズを可能にした

出版社: 東京図書

ISBN-10: 4489021712

ISBN-13: 978-4489021718

発売日: 2014/01/09

氏名：山川 義介

- 1981年 横浜国立大学工学部卒業
- 1981年 TDK株式会社入社  TDK
磁気テープ研究部～記録メディア事業部（8ミリビデオ、S-VHSビデオ開発）
- 1989年 商品企画課（カセットテープ、MD）
- 1992年 株式会社マルマン入社  maruman
常務取締役家電事業部長兼マーケティング部長
- 1995年 株式会社エムアンドシー設立  M&C
インターネット通販
インターネットリサーチシステム構築
- 2000年 株式会社インタースコープ設立  InterScope®
インターネットリサーチ&マーケティングコンサルティング
- 2005年 株式会社ALBERT設立  Albert
Recommendation technology
レコメンデーションの専門企業（レコメンドエンジン、CRMソリューション）
- 2008年 明治大学大学院グローバル・ビジネス研究科（MBA）非常勤講師

詳細はWikipediaをご参照ください。