

# 金銭的価値による商品ランキングのための併買関係分析

横田祐介\*<sup>1</sup> 松井藤五郎\*<sup>2</sup> 大和田勇人\*<sup>2</sup>  
Yusuke Yokota Tohgoroh Matsui Hayato Ohwada

\*<sup>1</sup>東京理科大学大学院 理工学研究科 経営工学専攻

Department of Industrial Administration, Graduate School of Science and Technology, Tokyo University of Science

\*<sup>2</sup>東京理科大学 理工学部 経営工学科

Department of Industrial Administration, Faculty of Science and Technology, Tokyo University of Science

We propose a method for making item ranking with the pecuniary values by analyzing cross-bought items based on association rules. As a result, we can find the item that was simultaneously bought with a lot of the other items and pecuniary value is high. We also show the experimental results with FoodMart data and Musashi project data, which indicates the proposed method is efficient.

## 1. 序論

現在、アソシエーションルールを用いたバスケット分析から商品の併買関係を明らかにし、販売戦略に生かす研究が多数なされている。バスケット分析とは、スーパーマーケットなどの顧客の買い物籠の内容を分析することで、販売促進のためのセール商品や店舗レイアウトなど様々な局面で有用することができる [1]。また、一般的なデータ解析に適用できる柔軟なものであるため、データマイニングの主要技術の一つとして位置づけられている。実際には支持度や確信度という指標を用いて商品間の関係性を明らかにしている。

近年ではアソシエーションルールを用いて併売関係を利用した商品の様々な研究が行われている。併売は、商品を販売する時に、複数の商品をセットにして意図的に販売するものである。しかし、本研究では併売ではなく併買に着目して研究を行った。ここでいう併買とは、ある商品  $X$  が購入されたときに他の商品が意図的に同時購入されていることを指す。自分が買われることで他の商品を同時に購入させる確率が高い商品は金銭的価値が高いと言える。

関連研究として八木 [2] らは、アソシエーションルールとページランクから各アイテムの重要度の分析を行って、併買しやすい商品のランキングを作成している。TOM BRJIS [3] らは、アソシエーションルールを用いて商品を販売するのに最適な選択を行うためのモデルを構築している。奥津 [4] らは、利益効率性という指標を用いてアソシエーションルールを評価するシステムの構築を行っている。

この他にも併買関係に注目して商品の価値を測る様々な研究が存在しているが、商品の併買関係のみで販売利益を考慮できていないことから、実際に価値のある商品を判断しにくくなっている。

そこで本論文では、金銭的価値による商品のランキングのために商品の併買関係を分析し、販売利益も考慮したランキングを作成する。ランキングにすることによって、たくさんの商品を併買していて金銭的価値のある商品を見つけることができる。

## 2. 提案手法

商品の売上総利益から経費を引いたものを販売総利益、併買関係のあるそれぞれの商品から分配される利益を併買利益とする。商品自体の販売総利益が高いことかつ販売総利益が高い多数の商品から併買されている商品は金銭的価値が高い商品として計算手法を提案する。今回は併買利益の配分方法としてアソシエーションルールの確信度 [5] を利用する。アソシエーションルール  $X \rightarrow Y$  の確信度  $conf(X \rightarrow Y)$  はすべてのトランザクション (一人の顧客が一回の買い物で購入したアイテムの集合) の中で  $X$  を含むトランザクションのうち、 $X$  と  $Y$  を共に含むトランザクションの割合で計算される。

### 2.1 金銭的価値の計算方法

利益の受け渡しの計算は二つの操作からなっており、それを操作  $(i)$ ,  $(ii)$  とする。操作  $(i)$  では販売総利益からの併買利益の分配を、操作  $(ii)$  では他の商品からくる併買利益からの併買利益の分配を行っている。

操作  $(i)$  販売総利益  $I_0$  の一部をユーザが設定した併買寄与率  $p$  の割合によって他の商品への併買利益とし、併買関係のある商品に分配する。この  $p$  の値は各商品によってかわるものとし、 $p_1, p_2, \dots, p_m$  とする。

次に、併買関係のある商品に併買利益の分配を行い、残りがそれぞれの販売総利益からの金銭的価値の値となる。つまり、他の商品に与える併買利益  $O_1, O_2, \dots, O_m$  の計算は、 $O_m = p_m I_0$  となる。

販売総利益から他の商品に与える併買利益を引いて残った値を  $R_0$  とすると、 $R_0 = I_0 - \sum_{i=1}^m p_i I_0$  となる。

操作  $(ii)$  それぞれの商品には併買関係のある商品からの併買利益がやってくる。併買関係のある商品からきた併買利益をそれぞれ  $I_1, I_2, \dots, I_n$  とする。それらを合計した値  $\sum_{i=1}^n I_i$  は、商品  $X$  が併買関係のある商品に与えていた併買利益である。

この併買利益の一部を併買寄与率  $p$  によって併買関係のある商品への併買利益とみなし、再び受け渡しを行う。つまり、他の商品に与える併買利益  $O_1, O_2, \dots, O_m$  の計算は、 $O_m = p_m \sum_{i=1}^n I_i$  となる。

併買関係のある商品に併買利益の分配を行い、残りがそれぞれの併買利益からの金銭的価値の値となる。他の商品か

連絡先: 横田祐介, 東京理科大学大学院 理工学研究科  
経営工学専攻, 千葉県野田市山崎 2641, 04(7124)1501,  
j7408635@ed.noda.tus.ac.jp

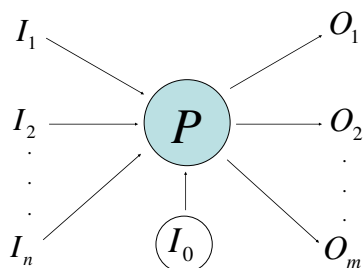


図 1: 商品 X と併買利益の関係

らきた併買利益から他の商品に与える併買利益を引いて残った値を  $R_s$  とすると、 $R_s = \sum_{i=1}^n I_i - \sum_{j=1}^m p_j \sum_{i=1}^n I_i$  となる。ここで、 $s$  は操作 (ii) が実行されている回数である。一番最初は  $R_1$  となる。一つ前までの金銭的価値に操作 (ii) で出た金銭的価値を足す。

三回目以降は操作 (ii) の繰返しである。操作を繰り返すごとに併買利益の一部が商品の金銭的価値として固定されていく。循環している併買利益の合計が、ユーザが定義した  $w$  の値を下回ったときに計算が終了する。操作 (ii) が  $s$  回行われて計算が終了したと考えると、金銭的価値  $P$  は、 $P = \sum_{i=0}^s R_i$  となる。

## 2.2 金銭的価値と併買寄与率 $p$ の定義

金銭的価値と今回の実験に用いる併買寄与率  $p$  の定義を以下に示す。

1. 各商品は固有の販売総利益を持っている。
  2. ある商品  $X$  に対して
    - ・  $X$  の販売利益を  $I_0$  とする。
    - ・他の商品から  $X$  へくる各々の併買利益を  $I_1, I_2, \dots, I_n$  とする。
    - ・  $X$  から他の商品に与える各々の併買利益を  $O_1, O_2, \dots, O_m$  とする。
    - ・  $X$  が他の商品に与える利益の割合値を併買寄与率  $p$  とする。
    - ・  $p$  の値はユーザが自由に設定できるものであり、商品ごとにその値は異なる。
    - ・  $X$  の金銭的価値を  $P$  とする。
- このとき、次が成り立つものとする。

$$P = (1-p) \sum_{i=0}^n I_i, \quad \sum_{j=1}^m O_j = p \sum_{i=0}^n I_i$$

今回は併買寄与率  $p$  の値を、確信度の値を基本として販売個数の割合で正規化を行い、次のようにして決めた。

- ・ 商品  $X$  から併買利益が与えられる商品の個数を  $m$  とする。
  - ・ 商品  $X$  から利益が与えられる各々の商品を  $X_1, X_2, \dots, X_m$  とする。
  - ・ 商品  $X$  から与える併買利益の割合を  $p$  とする。
  - ・ 各商品へ与える併買利益の割合を  $p_1, p_2, \dots, p_m$  とする。
  - ・ 商品の各々の売り上げ個数を  $w_1, w_2, \dots, w_m$  とする。
  - ・ アソシエーションルール  $A \rightarrow B$  の確信度を  $conf(A \rightarrow B)$  とする。
- このとき次が成り立つものとする。

$$p = \sum_{i=1}^m p_i, \quad O_i = p_i \sum_{j=0}^n I_j$$

表 1: 事前実験の結果

順位	元の順位	商品名	利益 (円)	元の利益 (円)
1	1	油揚げ 3	248,814.1	409,200
2	8	スバゲッティ 3	238,791.6	84,237
3	2	冷凍調理 5	206,849.5	243,775
4	3	紅茶 5	203,523.1	232,766
5	4	炭酸フレーバー 2	199,609.7	211,066
6	5	動物性油脂 2	195,823	193,307
7	6	ピーナッツ・チョコクリーム 1	194,081.3	177,850
8	7	デザート素 2	180,932.9	151,657
9	9	コーラ 2	48,728.6	13,296

$$p_i = \frac{w_i}{\sum_{j=1}^m w_j} \times conf(X_i \rightarrow X)$$

商品  $X$  と併買関係にある商品との利益の関係は図 1 のようになる。

$p$  を決めることによって確信度と販売個数に基づいて利益を分配できるようにしたが、変動する金額が小さくなることを考慮して、ユーザがパラメータ ( $0 < 1$ ) の値を与えることで配分する利益の値を全体的に引き上げられるようにした。各商品の  $p$  を  $p_1, p_2, \dots, p_k$  とし、 $p = 0$  の時  $p_k$  の値は変わらず、 $p = 1$  の時  $p$  の合計は 1 となるように設定している。

$$p'_k = \frac{(1 - p) \times \sum_{i=1}^k p_i + p_k}{\sum_{i=1}^k p_i} \times p_k$$

## 3. 実験

事前実験、実験共に OS は Windows XP、ブラウザは Internet Explorer 7、この他に eclipse 3.2、MySQL 5.1、TOM-CAT 5.5 などを使い JAVA で実装した。大量のトランザクションと商品を持つデータで実験を行う前に、少量のトランザクションと商品を持つデータでの事前実験を行った。

### 3.1 Musashi project のデータを用いた事前実験

事前実験に用いられるデータは Musashi project が提供するサンプルのデータ、200201data.txt を用いた。計算対象のトランザクション数は 3,099 個、商品は 10 個である。

実験結果のランキングを表 1 に示した。実験結果の表示は順位、元の順位、商品名、利益 (円)、元の利益 (円) の 5 つを表示した。元の順位とは、それぞれの商品の販売総利益による順位、元の利益はその金額である。順位とは、提案した手法で計算したあとの金銭的価値の順位、利益はその金額である。

### 3.2 FoodMart による実験

Microsoft SQL Server Analysis Services のサンプルキューブである FoodMart を用いて実験を行った。FoodMart 中の sales1998dec 表のトランザクション 3781 個、商品 1559 個を用いて行った。すべての商品の併買関係を計算すると膨大な時間がかかるため、今回はアソシエーションルールの支持度を閾値として用いた。支持度が 0.004 と 0.005、 $\alpha = 0, 0.3, 0.5, 0.7$  の  $2 \times 4 = 8$  通りの実験を行った。なお、支持度 0.004 のときは 1,559 個中 222 個の商品が、0.005 のときは 49 個の商品が計算対象となった。今回、金銭的価値の計算の終了条件は  $\epsilon = 1$  とし、事前の実験から 10,000 回繰返しせば条件を満たすことを確認して 10,000 回計算した。

表 2 に実験前の販売総利益によるランキング、表 3 に支持度 0.004、 $\alpha = 0.5$  のときの金銭的価値を出したランキングの 15 位までを例として示した。利益と元の利益の値が異なる商品が金銭的価値の値が計算された商品である。

表 2: 実験前の販売総利益によるランキング

順位	商品名	利益 (円)
1	ブッカー 低脂肪カッテージチーズ	18,370
2	ヒルトップ MG イブプロフェン	18,080
3	エヴァーベター チーズボール	17,270
4	山源 ドイツ産白ワイン	16,850
5	ベスト・チョイス チョコレートチップスクッキー	16,440
6	オランダ村 レーズンクッキー	16,220
7	ベスト・チョイス ポテトチップス	15,760
8	ヘルモンド バックサラダ	15,430
9	ぶたのしっぽ 冷凍コーン	15,170
10	オランダ村 クッキー	14,910
11	王様の冠 オレンジジャム	14,430
12	赤い翼 プラスティックナイフ	14,400
13	ヒルトップ ミントマウスウォッシュ	14,390
14	ハイマウンテン 新じゃが	14,290
15	メジャートップ ワイン(シャブリ)	14,150

表 3: 支持度 0.004、 $\alpha=0.5$  のときの実験結果の一部

順位	元の順位	商品名	利益 (円)	元の利益 (円)
1	2	ヒルトップ MG イブプロフェン	14,960.49	18,080
2	4	山源 ドイツ産白ワイン	14,697.22	16,850
3	1	ブッカー 低脂肪カッテージチーズ	14,091.16	18,370
4	21	ブッカー ストロベリーヨーグルト	13,660	13,660
5	3	エヴァーベター チーズボール	13,642.37	17,270
6	24	ファビュラス ストロベリードリンク	13,570	13,570
7	25	スプリム 絹クリーナー	13,550	13,550
8	28	トリスデイト 赤ビーマン	13,270	13,270
9	29	ベアード 高級サーロインステーキ	13,240	13,240
10	22	Fort William's 乾燥ナツメヤシ	12,911.75	13,630
11	7	ベスト・チョイス ポテトチップス	12,898.93	15,760
12	36	カールヤ	12,590	12,590
13	8	ヘルモンド バックサラダ	12,410.59	15,430
14	10	オランダ村 クッキー	12,396.28	14,910
15	39	ホワイトピーチ ベーグル	12,390	12,390

#### 4. 考察

事前実験では、商品数が少ないため、抽出したアソシエーションルールの確信度の値が 0.9 から 1 の間の数値が大半を占めた。また、スパゲッティ3 が大きく金銭的価値の値を上げていた。これは、抽出したアソシエーションルールから他の商品に対してたくさんの商品を併買しているが、自分は他の商品から併買される確率が低く、他の商品からたくさんの利益をもらったと考えられる。確信度の値が、他の商品を併買するときは 0.9 から 1 の値を取っていたが、他の商品に併買されるときは 0.6 前後の値が多いことが結果に出ていると考えられる。ランキングは確信度の値に基づいて利益の配分が行われたことから併買関係を考慮したランキングができたと考えられる。

FoodMart のデータベースでは、トランザクションの傾向として商品数が 3 から 6 までに収まるものが大半で、確信度の値で大きな値を出しづらいつと考えられる。ただし、 $\alpha$  の値を用いて併買寄与率  $p$  の値を引き上げることができるので大きな問題ではないと考えられる。

実験では今回提案した併買寄与率  $p$  は、アソシエーションルールの確信度が低くなることから全体として小さな値をとることがわかった。計算手法として併買利益が大きく移動するのは一回目だけになることが予想されるが、0.001 ~ 0.004 の値が多く、 $\alpha$  が 0 のときは併買利益の大きな変動はなかった。

また、 $\alpha$  を 0.9 にして計算を行うと、全体に循環する利益が大きくなって一回の計算で固定される金銭的価値の値が大幅に減る。つまり計算を繰り返しても金銭的価値の値が決定できなくなると考えられる。

$\alpha$  の値を用いなければ大幅なランキング変動はないが、大量のデータがあるデータベースでも併買関係に基づいたランキングは作成されたことがわかった。

#### 5. 結論

本研究では、併買関係の分析から金銭的価値による商品のランキングを行った。併買利益の計算は繰返し計算を用いた手法を提案し、商品ごとに販売利益と併買利益からなる商品の金銭的価値を定義して計算できるようにした。

事前実験の結果から提案手法の実装が行われていることを確認して FoodMart により実験を行った。実験結果から併買関係に基づいた金銭的価値のランキングができたと考えられる。アソシエーションルールを用いて併買関係を考慮したランキングを作成することは、販売戦略を立てる上でも有効であると考えられる。

今後の展望として、様々なデータベースに使用できるシステムの設計、実際の販売店からの販売データを用いての実験などが考えられる。

#### 参考文献

- [1] 岡田孝, 元田浩. 相関ルールとその周辺. オペレーションズ・リサーチ, Vol. 47, .
- [2] 八木一光, 岡野慎吾, 森本康彦. 相関ルールの 2 部グラフを用いた重要アイテムの発掘. DEWS2006 6A-i7, 2006.
- [3] TOM BRIJIS, GILBERT SWINNEN, KOEN VAN-HOOF, and GEERT WETS. Building an association rules framework to improve product assortment decisions. *Data Mining and Knowledge Discovery*, Vol. 8, pp. 7-23, 2004.
- [4] 奥津元靖, 松井藤五郎, 大和田勇人. 東京理科大学 理工学部 経営工学科 平成 18 年度修士論文. 2007.
- [5] 喜連川優. データマイニングにおける相関ルール抽出技法. 人工知能学会誌, Vol. 12, .