日本オペレーションズ・リサーチ学会
2022年 秋季研究発表会

最適決定木を用いた処方的価格最適化

05001584 株式会社リクルート *池田 春之介 IKEDA Shunnosuke
非会員 株式会社リクルート 西村 直樹 NISHIMURA Naoki
02203940 筑波大学 鮏川 矩義 SUKEGAWA Noriyoshi
01309510 筑波大学 高野 祐一 TAKANO Yuichi

1. はじめに

価格は商品の需要に大きく影響し，適切に価格を設定することで，企業の収益や利益を最大化することができる．近年では，情報通信技術の発展により，需要の情報を適時反映することが可能になり，ますます価格決定の収益や利益に与える影響が大きくなっているため，その重要性は高まっている．

従来の研究では，単一商品の価格とその需要の関係性に着目するものが多かったが，回帰技術の進歩により，数多くの商品の需要モデリングが可能になり，複数商品の交差価格弾力性を考慮した需要モデリングも注目されている．先行研究 [1, 2] では，複数商品の価格弾力性を非線形回帰で予測し，その回帰曲線に基づいた収益最大化問題を 0-1 整数二次最適化（BQO）問題として定式化し，高速なアルゴリズムを提案している．これらの研究では陽にモデルを記述できる非線形回帰を用いているが，価格弾力性は時期や商品などの購買条件の変化で多様な曲線を描くことが知られており，形状に制約のない回帰モデルの使用が望ましい [3]．

また，価格最適化において，価格と需要の関係性を正確にモデリングすることは収益の向上に直結するため，より精度の高い回帰モデルを用いることが求められる．LightGBM や XGBoost といった勾配ブースティング木による手法が多くのデータセットにおいて高い汎化性能を実現することが知られている．ただし，これら勾配ブースティング木系の回帰モデルでは，求解可能な価格最適化問題を定式化することが難しく，最適な価格を求めるためには価格の組合せを列挙して全探索する必要がある．そのため，数十個の商品でも現実的な時間で解くことは困難となってしまう．

そこで本研究では，高精度の回帰モデルを用いつつ，現実的な時間で最適解が求められる価格最適化モデルを提案する．具体的には，回帰モデルに解釈性を失わずに高い汎化性能が期待できる最適決定木 [4] を用いる．それにより，価格最適化問題を混合整数線形最適化（MILO）問題として定式化することができ，求解可能な最適化問題として記述できる．

2. 問題設定

本研究では，収益最大化を目的とした複数商品の価格最適化問題を扱う．

各商品 $m \in \mathcal{M}$ の価格と需要をそれぞれ p_m，$q_m(\mathbf{p})$ とする．ただし，$\mathbf{p} = (p_m)_{m\in\mathcal{M}}$ である．商品 m の需要 $q_m(\mathbf{p})$ は，自身と他の商品の価格に応じて決まるため，価格ベクトル $\mathbf{p}$ の関数として記述される．また，価格 p_m は $|\mathcal{K}|$ 個の価格候補の集合 $\{P_{mk} \mid k \in \mathcal{K}\}$ から選ばれるものとする．これは，5%オフ，10%オフや 100 円引き，1000 円引きなどの実際に行われている価格設定に対応し，より現実に即していると言える．

これらを踏まえて，複数商品の価格最適化問題は一般的に次のように定式化できる．

$$\max_{\mathbf{p}} \quad \sum_{m\in\mathcal{M}} p_m q_m(\mathbf{p}) \tag{1}$$

$$\text{s.t.} \quad p_m \in \{P_{mk} \mid k \in \mathcal{K}\}, \quad \forall m \in \mathcal{M} \tag{2}$$

先行研究 [1, 2] では，需要 $q_m(\mathbf{p})$ を非線形回帰モデルで表現し，BQO 問題として定式化している．また，需要 $q_m(\mathbf{p})$ に対して，ブラックボックスな回帰モデルを用いた場合は，解くのに価格の組合せを $|\mathcal{K}|^{|\mathcal{M}|}$ 通り列挙し，全探索する必要がある．

3. 最適決定木を用いた定式化

本節では，需要 q_m の予測に最適決定木 [4] を用いた場合の価格最適化問題を定式化する．なお，最適決定木の構築は，最適化問題を解く前に行うことができるため，最適決定木の構造は所与の定数として扱うことができる．以下，使用する記号の定義を示す．

集合

$\mathcal{D}$ 価格以外の説明変数の添字集合
$\mathcal{M}$ 商品の集合
$\mathcal{R}_t/\mathcal{L}_t$ 葉ノード t へ向けて右/左に分岐する枝ノードの集合
$\mathcal{T}_B/\mathcal{T}_L$ 枝/葉ノードの集合

定数

β_{jt} 葉ノード t での j 番目の説明変数の回帰係数
$\varepsilon/\varepsilon_{\max}$ 説明変数の値の最小間隔/最小間隔の最大値

a_{dt}　ノード t において説明変数 d で分岐する場合に1，そうでない場合に0となる定数

b_t　ノード t での境界を表す定数

g_d　価格以外の d 番目の説明変数

M_q　十分大きな正定数

P_{mk}　商品 m の k 番目の価格候補

決定変数

p_m　商品 m の価格を表す連続変数

q_m　商品 m の需要を表す連続変数

x_{mk}　商品 m が k 番目の価格をとる場合に1，そうでない場合に0となる0-1変数

z_{mt}　商品 m が葉ノード t に割り当てられる場合に1，そうでない場合に0となる0-1変数

商品 m ごとに最適決定木により需要が予測されるため，最適決定木に関する記号は，各商品 m ごとに定義される．表記上，それらの記号は上付きの (m) として表現している．定式化を以下に示す．

$$\max_{\mathbf{p},\mathbf{q},\mathbf{x},\mathbf{z}} \quad \sum_{m\in\mathcal{M}} p_m q_m \tag{3}$$

s. t.

$$q_m - \left(\sum_{m'\in\mathcal{M}} \beta^{(m)}_{m't} p_{m'} + \sum_{d\in\mathcal{D}^{(m)}} \beta^{(m)}_{dt} g_d\right) \geq -M_q^{(m)}(1-z_{mt}), \quad \forall t\in\mathcal{T}_L^{(m)}, \forall m\in\mathcal{M} \tag{4}$$

$$q_m - \left(\sum_{m'\in\mathcal{M}} \beta^{(m)}_{m't} p_{m'} + \sum_{d\in\mathcal{D}^{(m)}} \beta^{(m)}_{dt} g_d\right) \leq M_q^{(m)}(1-z_{mt}), \quad \forall t\in\mathcal{T}_L^{(m)}, \forall m\in\mathcal{M} \tag{5}$$

$$\sum_{m'\in\mathcal{M}} a^{(m)}_{m't'} p_{m'} + \sum_{d\in\mathcal{D}^{(m)}} a^{(m)}_{dt'} g_d, \geq b^{(m)}_{t'} - (1-z_{mt}), \quad \forall t'\in\mathcal{R}_t^{(m)}, \forall t\in\mathcal{T}_L^{(m)}, \forall m\in\mathcal{M} \tag{6}$$

$$\sum_{m'\in\mathcal{M}} a^{(m)}_{m't'} p_{m'} + \sum_{d\in\mathcal{D}^{(m)}} a^{(m)}_{dt'} g_d + \epsilon^{(m)} \leq b^{(m)}_{t'} + (1+\epsilon_{\max})(1-z_{mt}), \quad \forall t'\in\mathcal{L}_t^{(m)}, \forall t\in\mathcal{T}_L^{(m)}, \forall m\in\mathcal{M} \tag{7}$$

$$p_m = \sum_{k\in\mathcal{K}} P_{mk} x_{mk}, \quad \forall m\in\mathcal{M} \tag{8}$$

$$\sum_{k\in\mathcal{K}} x_{mk} = 1, \quad \forall m\in\mathcal{M} \tag{9}$$

$$\sum_{t\in\mathcal{T}_L^{(m)}} z_{mt} = 1, \quad \forall m\in\mathcal{M} \tag{10}$$

$$x_{mk}, z_{mt} \in \{0,1\}, \quad \forall t\in\mathcal{T}_L^{(m)}, \forall m\in\mathcal{M}, \forall k\in\mathcal{K} \tag{11}$$

目的関数 (3) は，総収益の最大化を表す．式 (4),(5) は，各商品の需要 q_m が葉ノード t に割り当てられた際に，その葉ノードでの回帰モデルの適用を意味する．式 (6),(7) は，各商品の需要 q_m が葉ノード t に割り当てられる場合に、葉ノード t への分岐条件をすべて満たしていることを表す．式 (8),(9) は，商品 m の価格は価格候補から一つだけ選ばれることを表す．また，式 (10) にて，各商品の需要 q_m は必ずどれか一つの葉ノードに割り当てられる．最後に式 (11) で，変数 x_{mk}, z_{mt} を0-1変数としている．

本定式化は，目的関数が変数の積となっており，非線形の混合整数最適化問題となっているため，商用ソルバーでも厳密解を求めるのは非常に難しい．そこで，目的関数 (3) に式 (8) を代入し，新たに $u_{mk} = x_{mk}q_m$ となる変数を導入することで，目的関数を線形化することができる．加えて，次のような制約を追加することで，元の定式化はMILO問題に等価変換できる．

$$-M_q^{(m)} x_{mk} \leq u_{mk} \leq M_q^{(m)} x_{mk}, \quad \forall m\in\mathcal{M}, \forall k\in\mathcal{K} \tag{12}$$

$$q_m - M_q^{(m)}(1-x_{mk}) \leq u_{mk} \leq q_m + M_q^{(m)}(1-x_{mk}), \quad \forall m\in\mathcal{M}, \forall k\in\mathcal{K} \tag{13}$$

MILO問題に変換した定式化を以下に示す．

$$\max_{\mathbf{p},\mathbf{q},\mathbf{u},\mathbf{x},\mathbf{z}} \quad \sum_{m\in\mathcal{M}} \sum_{k\in\mathcal{K}} P_{mk} u_{mk} \tag{14}$$

s. t.　式 (4)–(13)

4. 数値実験

提案手法の有効性を検証するために，既存手法との比較実験を実施する．実験結果の詳細は当日報告する．

参考文献

[1] Ito, Shinji, and Ryohei Fujimaki. “Large-scale price optimization via network flow.” Advances in Neural Information Processing Systems, 29 (2016).

[2] Ito, Shinji, and Ryohei Fujimaki. “Optimization beyond prediction: Prescriptive price optimization.” Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2017.

[3] 杉田善弘, 上田隆穂, 守口剛. “プライシング・サイエンス―価格の不思議を探る―.” 同文館出版, 2005.

[4] Dunn, Jack William. “Optimal trees for prediction and prescription.” Ph.D. Thesis, Massachusetts Institute of Technology, 2018.