F-023

予測評価値の不確実性を考慮した推薦商品集合のロバスト最適化

柳 智也 †　池田 春之介 ‡　高野 祐一 §
筑波大学 †　筑波大学 ‡　筑波大学 §

1 はじめに

推薦システムとは，過去に利用者が商品に与えた評価をもとに，その利用者が好むであろう商品を提示する技術である．推薦システムでは，評価値の予測精度を高める研究が盛んに行われてきたが，近年は推薦する商品の多様性や新規性といった指標も重要視されている．

多様性を考慮した推薦手法として，平均・分散モデルが提案されている [1]．平均・分散モデルは金融商品のポートフォリオの構築に利用されてきたが，推薦システムに活用することで，同一商品の評価値のばらつきや商品間の評価値の関係性を考慮して推薦商品を決定することができる．また，平均・分散モデルで使用する共分散行列について，目標行列への縮小推定によって推定精度を向上させる方法が提案されている [2]．

しかし，先行研究 [1, 2] には推定統計量の不確実性を考慮していないという課題が存在する．平均・分散モデルでは商品の評価値の平均と共分散を推定する必要があり，これらは機械学習モデルと観測データを利用して推定される．このときに生じる推定誤差を無視して平均・分散モデルを利用した場合に，推薦システムの性能が低下する可能性がある．

本研究では，推定統計量の不確実性を考慮したロバスト平均・分散モデルを提案する．提案手法ではロバスト最適化 [3] を平均・分散モデルに適用することで，商品の評価値や共分散に推定誤差が生じる場合を考慮して推薦商品集合を決定することができる．

2 平均・分散モデル

本研究では，各利用者に推薦する商品の集合を決定する問題を扱う．μ_i は商品 i の予測評価値，σ_{ij} は商品 i, j の評価値の共分散を表し，これらは過去に利用者が商品に与えた評価値が格納された行列（評価値行列）を用いて推定を行う．また，n は推薦候補の商品数，N は推薦する商品数を表す．ここで，w_i を商品 i を推薦する場合に 1，そうでないときに 0 をとる 0-1 変数とすると，先行研究における平均・分散モデルは以下のような 0-1 整数最適化問題として定式化される．

$$\underset{\boldsymbol{w}}{\text{maximize}} \quad (1-\alpha)\sum_{i=1}^{n} w_i\mu_i - \alpha\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} \tag{1}$$

$$\text{subject to} \quad \sum_{i=1}^{n} w_i = N \tag{2}$$

$$w_i \in \{0,1\} \quad (i \in [n]) \tag{3}$$

式 (2) は利用者に N 個の商品を推薦する制約である．目的関数の第一項は予測評価値の合計を表しており，予測評価値が高い商品を推薦するほど大きな値を取る．一方で，第二項は合計予測評価値の分散を表しており，同一商品に対する評価値のばらつきや商品間の評価値の相関を考慮している．平均・分散モデルでは，これらの項をパラメータ α を用いて重み付けることで，予測評価値だけでなく商品間の関係性を考慮しながら推薦商品集合を決定する．結果と

Robust optimization of recommended items considering uncertainty of predicted ratings
† Tomoya Yanagi, University of Tsukuba
‡ Shunnosuke Ikeda, University of Tsukuba
§ Yuichi Takano, University of Tsukuba

して，推薦商品の多様性が向上し，利用者の満足度が高まることが期待できる．

一方で，評価値行列は疎であることと，機械学習モデルには推定誤差が生じることから，μ_i および σ_{ij} の不確実性を考慮しない場合に推薦の精度が低下する可能性がある．そのため，これらの誤差を考慮して推薦商品集合を決定することが望ましい．

3 ロバスト平均・分散モデル

本研究では，推定統計量の不確実性を考慮するためにロバスト最適化を適用する．ロバスト最適化では，不確実な数値の変動幅を設定し，最悪な状況が発生した場合を想定して最適化を行う．はじめに，平均と共分散の不確実性集合をそれぞれ $U_{\boldsymbol{\mu}}, U_{\boldsymbol{\Sigma}}$ で定義する．ここで，δ_i, ϵ_{ij} はそれぞれ平均と共分散の変動幅である．

$$U_\mu = \{\tilde{\boldsymbol{\mu}} \in \mathbb{R}^n \mid \mu_i - \delta_i \le \tilde{\mu}_i \le \mu_i + \delta_i, i \in [n]\}$$

$$U_\Sigma = \left\{\tilde{\boldsymbol{\Sigma}} \in \mathbb{R}^n \mid \sigma_{ij} - \epsilon_{ij} \le \tilde{\sigma}_{ij} \le \sigma_{ij} + \epsilon_{ij}, i, j \in [n]\right\}$$

また，平均が最悪ケースを取る商品数を Γ_μ，共分散が最悪ケースを取る商品の組数を Γ_Σ とする．このとき，ロバスト平均・分散モデルは以下のように定式化できる．

$$\begin{aligned} \underset{\boldsymbol{w}}{\text{maximize}} \quad & (1-\alpha)\left(\sum_{i=1}^n w_i \tilde{\mu}_i - A\right) \\ & - \alpha\left(\sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{\sigma}_{ij} + B\right) \quad (4) \\ \text{subject to} \quad & 式 (2), (3) \quad (5) \end{aligned}$$

ここで，A，B はそれぞれ以下の通りである．

$$A = \underset{\{S_\mu | S_\mu \subseteq [n], |S_\mu| \le \Gamma_\mu\}}{\text{maximize}} \sum_{i \in S_\mu} w_i \delta_i \quad (6)$$

$$B = \underset{\{S_\Sigma | S_\Sigma \subseteq [n]\times[n], |S_\Sigma| \le \Gamma_\Sigma\}}{\text{maximize}} \sum_{(i,j) \in S_\Sigma} w_i w_j \epsilon_{ij} \quad (7)$$

ここで，二つの最大化問題 (6)，(7) について双対を取ることで，次の単一の最適化問題に変形することができる．なお，z, g, p_i, q_{ij} は双対変数である．

$$\begin{aligned} \underset{\boldsymbol{w},\boldsymbol{p},\boldsymbol{q},z,g}{\text{maximize}} \quad & (1-\alpha)\left(\sum_{i=1}^n w_i \tilde{\mu}_i - z\Gamma_\mu - \sum_{i=1}^n p_i\right) \\ & - \alpha\left(\sum_{i=1}^n \sum_{j=1}^n w_i w_j \tilde{\sigma}_{ij} + g\Gamma_\Sigma + \sum_{i=1}^n \sum_{j=1}^n q_{ij}\right) \quad (8) \\ \text{subject to} \quad & 式 (2), (3) \quad (9) \\ & \delta_i w_i \le z + p_i \quad (i \in [n]) \quad (10) \\ & p_i \ge 0 \quad (i \in [n]) \quad (11) \\ & z \ge 0 \quad (12) \\ & w_i w_j \varepsilon_{ij} \le g + q_{ij} \quad (i, j \in [n]) \quad (13) \\ & q_{ij} \ge 0 \quad (i, j \in [n]) \quad (14) \\ & g \ge 0 \quad (15) \end{aligned}$$

ここで，式 (8)，(13) における 0-1 変数の積 $w_i w_j$ は線形化できる [4]．結果として，単一の混合整数線形最適化問題を解くことで各利用者に推薦する商品集合を決定することができる．

4 数値実験

提案手法の性能を推薦の精度および推薦商品の多様性の 2 つの指標で検証する．実験では，平均が最悪ケースを取る商品数 Γ_μ および共分散が最悪ケースを取る商品の組数 Γ_Σ を変化させた場合に推薦の精度と多様性がどのように変化するかを検証する．実験に使用したデータおよび結果の詳細は当日報告する．

参考文献

[1] Hurley, N.， Zhang, M. (2011). Novelty and diversity in top-N recommendation—Analysis and evaluation. ACM Transactions on Internet Technology (TOIT), 10(4), 1-30.

[2] Yasumoto, Y.， Takano, Y. (2023). Mean-variance portfolio optimization with shrinkage estimation for recommender systems. Optimization Online.

[3] Bertsimas, D.， Sim, M. (2004). The price of robustness. Operations Research, 52(1), 35-53.

[4] Williams, H. P. (2013). Model Building in Mathematical Programming. John Wiley & Sons.