Multi-Metric Optimization: Fast and Flexible Re-ranking Over Multiple Business Metrics

Kiana Liaghat*, Mohammad Forouhesh†
* Machine Learning Lead, Jabama, Tehran, Iran
Email: k.liaghat@jabama.com
†Data Science Manager, Jabama, Tehran, Iran
Email: m.forouhesh@jabama.com.com

***Abstract*—Recent advances in recommendation systems have highlighted the limitations of single-metric optimization approaches, particularly in complex business environments where multiple objectives must be balanced simultaneously. This paper addresses the critical challenge of optimizing accommodation ranking systems that must satisfy multiple business metrics concurrently. We aim to develop a mathematically rigorous framework that enables joint optimization of multiple Key Performance Indicators (KPIs) while maintaining system stability and performance. Using a novel approach based on the Rayleigh quotient and generalized eigenvalue problems, we propose a mathematically rigorous solution that maximizes the aggregate squared correlation between ranking functions and multiple business metrics.**

***Index Terms*—Ranking, Optimization, Multi-criterion Optimization**

I. Introduction

Ranking, re-ranking, and recommendation systems are foundational in online platforms, directly influencing both user satisfaction and key business outcomes such as conversion rate (CR), profit, and retention. While traditional approaches often optimize a single objective—typically relevance or user engagement—in practice, platforms must balance multiple, sometimes competing, business metrics [1]–[3]. This has led to the development of multi-objective optimization (MOO) frameworks, integrating explicit business objectives or constraints alongside user-centric goals [4], [5].

A wide variety of MOO methods have been proposed, ranging from primal-dual and Lagrangian techniques with global business constraints [1], [3], [4] to direct label aggregation and model distillation for optimizing multiple objectives in learning-to-rank settings [5]–[7]. However, these methods often have key limitations: scalarization approaches may struggle with non-convex trade-offs [5], [7], and many industrial systems restrict flexibility by optimizing over a small number of objectives or by hard-coding trade-off weights [1], [3]. Comprehensive algorithms that provide theoretical optimality, support arbitrarily many business metrics, and allow explicit, flexible weighting of metrics according to business need remain rare.

In this work, we propose a simple, mathematically rigorous multi-metric optimization algorithm that addresses these limitations. The proposed algorithm is provably optimal, supports an arbitrary number of business and relevance metrics, and flexibly incorporates user-specified importance weights in case of need. This generality and theoretical guarantee fill an important gap in the MOO literature for ranking and recommendation systems.

II. Related Work

Explicit incorporation of business metrics as optimization objectives has been a major trend in ranking literature. Primal-dual and Lagrangian approaches have proven effective in real-time systems, using shadow prices derived from linear programming duality to enforce business constraints such as revenue or impressions [1], [3], [8]. These methods have demonstrated their scalability and practical value in large-scale production environments, including LinkedIn newsfeeds and Amazon's video homepage [2], [3].

Beyond constrained programming, recent work has explored treating multiple metrics as explicit joint objectives. Sun et al. [4] proposed online welfare maximization by optimizing all business requirements as objectives, providing adaptability and theoretical guarantees. Label aggregation and distillation methods have targeted offline learning, with stochastic methods offering full Pareto-front coverage under proper distributional assumptions [5]. Distillation-based solutions [6] and multi-label learning-to-rank [7] combine several business and relevance metrics, but may involve expensive hyperparameter tuning, offer limited explicit user control over importance weights, or involve heuristic rather than theoretically optimal solutions.

Combinatorial and economic approaches have also been used to balance profit and user utility, via integer programming [9], dynamic programming [10], or auction-theoretic ranking mechanisms [11]–[15]. However, these models often target specific domains, trade-off only two objectives at a time [10], [11], or lack explicit direct support for arbitrary business-defined importance weights.

Many frameworks, particularly primal-dual and label aggregation methods, are practically limited to small numbers of objectives or fixed business constraints, and can become complex or inflexible in high-dimensional or rapidly changing business environments [1], [3], [5]. Common scalarization and aggregation techniques often require manual, sometimes non-transparent tuning of weights, and their impact on Pareto optimality and solution coverage can be hard to guarantee [5],

[7]. While several methods show empirical success or theoretical justification for narrower problem classes, a simple, general-purpose algorithm with rigorous optimality guarantees for arbitrary, user-weighted business metrics remains absent in the current literature. Algorithms that can directly accept business stakeholder input (e.g., explicit importance values over multiple metrics) are rare, limiting the practical usability of many academic approaches.

The algorithm introduced in this work directly addresses these limitations. It supports optimization over any number of business and user-centric metrics relevant to the platform, allows explicit business-specified importance weights for each metric, and is mathematically proven to return the optimal solution for the specified metric importance vector, regardless of the number of objectives. It is computationally simple and directly usable in both industrial and academic settings. By enabling theoretically guaranteed, flexible, and stakeholder-aligned optimization across an arbitrary set of business metrics, our approach bridges a key gap in multi-objective ranking—one that current state-of-the-art methods address only partially [1], [3]–[7]. Our work thus provides a robust foundation for business-aware ranking optimization in modern, metric-rich environments.

III. Methodology

In this section, we introduce our proposed multi-metric optimization algorithm, provide necessary mathematical preliminaries, and present a formal proof of its optimality.

A. Mathematical Preliminaries

Definition 1 (Centered Variables)**.** *Assume $\boldsymbol{x}$ is a stochastic vector. Then $\boldsymbol{x}$ is centered if*

$$\tilde{\boldsymbol{x}} = \boldsymbol{x} - \bar{\boldsymbol{x}}, \quad \text{where} \quad \bar{\boldsymbol{x}} = \frac{1}{m}\sum_{i=1}^{m} x_i.$$

In the literature, the centered variable of $\boldsymbol{x}$ is denoted as $\tilde{\boldsymbol{x}}$. In this document, we omit this tilde and use the same notation after centering.

Definition 2 (Correlation)**.** *Assume $\boldsymbol{x}, \boldsymbol{y}$ are centered variables. Then, the **correlation** between $\boldsymbol{x}$ and $\boldsymbol{y}$ is defined as:*

$$\mathit{corr}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\mathit{cov}(\boldsymbol{x}, \boldsymbol{y})}{\sqrt{\mathit{var}(\boldsymbol{x}) \cdot \mathit{var}(\boldsymbol{y})}} = \frac{\boldsymbol{x}^\top \boldsymbol{y}}{\sqrt{(\boldsymbol{x}^\top \boldsymbol{x})(\boldsymbol{y}^\top \boldsymbol{y})}}. \tag{1}$$

*Correlation measures the **linear relationship** between two variables [16].*

Definition 3 (Rayleigh Quotient)**.** *Given symmetric matrices $\boldsymbol{A}, \boldsymbol{B}$, the Rayleigh quotient is defined as:*

$$\mathcal{R}(\boldsymbol{a}) = \frac{\boldsymbol{a}^\top \boldsymbol{A} \boldsymbol{a}}{\boldsymbol{a}^\top \boldsymbol{B} \boldsymbol{a}}. \tag{2}$$

Theorem 1 (Courant-Fischer)**.** *For symmetric $\boldsymbol{A}$ and positive definite $\boldsymbol{B}$, the maximum value of $\mathcal{R}(\boldsymbol{a})$ is the **largest generalized eigenvalue** of the equation $\boldsymbol{A}\boldsymbol{a} = \lambda \boldsymbol{B}\boldsymbol{a}$, achieved by the corresponding eigenvector [17].*

Proof. Assume $\boldsymbol{A}, \boldsymbol{B}$ are arbitrary symmetric matrices, $\boldsymbol{a}$ is an arbitrary vector, and $\mathcal{R}(\boldsymbol{a})$ is the Rayleigh quotient. The critical points occur where the gradient of $\mathcal{R}(\boldsymbol{a})$ with respect to $\boldsymbol{a}$ is zero:

$$\nabla_{\boldsymbol{a}} \mathcal{R}(\boldsymbol{a}) = \frac{2\boldsymbol{A}\boldsymbol{a}(\boldsymbol{a}^\top \boldsymbol{B}\boldsymbol{a}) - 2\boldsymbol{B}\boldsymbol{a}(\boldsymbol{a}^\top \boldsymbol{A}\boldsymbol{a})}{(\boldsymbol{a}^\top \boldsymbol{B}\boldsymbol{a})^2} = \boldsymbol{0}. \tag{3}$$

Now, canceling common terms and rearranging:

$$\boldsymbol{A}\boldsymbol{a}(\boldsymbol{a}^\top \boldsymbol{B}\boldsymbol{a}) = \boldsymbol{B}\boldsymbol{a}(\boldsymbol{a}^\top \boldsymbol{A}\boldsymbol{a}). \tag{4}$$

Let $\lambda = \frac{\boldsymbol{a}^\top \boldsymbol{A}\boldsymbol{a}}{\boldsymbol{a}^\top \boldsymbol{B}\boldsymbol{a}}$, which is the Rayleigh quotient itself. Substituting λ:

$$\boldsymbol{A}\boldsymbol{a} = \lambda \boldsymbol{B}\boldsymbol{a}. \tag{5}$$

This is the **generalized eigenvalue problem**. Thus, the Rayleigh quotient $\mathcal{R}(\boldsymbol{a})$ equals λ when $\boldsymbol{a}$ satisfies $\boldsymbol{A}\boldsymbol{a} = \lambda \boldsymbol{B}\boldsymbol{a}$. To maximize $\mathcal{R}(\boldsymbol{a})$, we seek the largest λ. □

B. Proposed Algorithm

We aim to find a linear function $f(\mathbf{z}) = \mathbf{z}\dashv^\top + b$ that maximizes the aggregate squared correlation with a set of k KPIs stored in a matrix $\mathbf{M} \in \mathbb{R}^{m\times k}$. [Multi Metric Optimization] Given $\mathbf{Z} \in \mathbb{R}^{m\times n}$ and $\mathbf{M} \in \mathbb{R}^{m\times k}$, and a linear operator $\Phi : \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times 1}$ defined as

$$\Phi(\mathbf{Z}) = \mathbf{Z}\dashv + b, \quad \dashv, b \in \mathbb{R}^{n\times 1},$$

find $\dashv$ to maximize the aggregate squared correlation between $\Phi(\mathbf{Z})$ and each column $\mathbf{M}_i$ of $\mathbf{M}$:

$$\max_{\Phi} \sum_{i=1}^{k} \mathrm{corr}^2(\Phi(\mathbf{Z}), \mathbf{M}_i). \tag{6}$$

Summing squared correlations across all KPIs in Equation (6) is a valid approach to aggregate the relationships between $\Phi(\mathbf{Z})$ and the KPIs into a single objective. This avoids cancellations between positive and negative correlations and ensures that $\Phi(\mathbf{Z})$ aligns collectively with the KPIs [**?**].

Algorithm 3.1: Multi-Metric Optimization

Using the properties of centered variables (see Definition 1), we can center both $\mathbf{Z}$ and $\mathbf{M}$. Since correlation is invariant to shifts and Φ is a linear operator, we define:

$$\tilde{\mathbf{Z}} = \mathbf{Z} - \mathbf{1}\bar{\mathbf{Z}}, \quad \tilde{\mathbf{M}} = \mathbf{M} - \mathbf{1}\bar{\mathbf{M}}, \tag{7}$$

where $\bar{\mathbf{Z}}, \bar{\mathbf{M}}$ are column means. The intercept b becomes zero, so we have $\Phi(\mathbf{Z}) = \mathbf{Z}\dashv$, and Equation (6) simplifies to:

$$\max_{\dashv} \sum_{i=1}^{k} \mathrm{corr}^2(\mathbf{Z}\dashv, \mathbf{M}_i). \tag{8}$$

Using the definition of correlation (see Definition 2), we expand each squared term:

$$\mathrm{corr}^2(\mathbf{Z}\dashv, \mathbf{M}_i) = \frac{(\dashv^\top \mathbf{Z}^\top \mathbf{M}_i)^2}{(\dashv^\top \mathbf{Z}^\top \mathbf{Z}\dashv)(\mathbf{M}_i^\top \mathbf{M}_i)}. \tag{9}$$

Summing over all k KPIs:

$$\sum_{i=1}^{k} \mathrm{corr}^2(\mathbf{Z}\dashv, \mathbf{M}_i) = \frac{\dashv^\top \mathbf{Z}^\top \mathbf{M}\mathbf{M}^\top \mathbf{Z}\dashv}{\dashv^\top \mathbf{Z}^\top \mathbf{Z}\dashv}. \tag{10}$$

This is a Rayleigh quotient (see Definition 3) of the form:

$$\mathcal{R}(\dashv) = \frac{\dashv^\top \mathbf{C}_1 \dashv}{\dashv^\top \mathbf{C}_2 \dashv}, \quad \text{where} \quad \mathbf{C}_1 = \mathbf{Z}^\top \mathbf{M}\mathbf{M}^\top \mathbf{Z}, \quad \mathbf{C}_2 = \mathbf{Z}^\top \mathbf{Z}. \tag{11}$$

According to the Courant–Fischer theorem (Theorem 1), the optimal solution to Equation (10) is obtained by solving the generalized eigenvalue problem (GEP):

$$\mathbf{C}_1 \dashv = \lambda \mathbf{C}_2 \dashv,$$

and choosing the eigenvector $\dashv$ corresponding to the largest eigenvalue λ.

Proof. We only need to verify Equation (10) and that the matrices in Equation (11) satisfy the conditions of Theorem 1.

First, assume that each KPI $\mathbf{M}_i$ is standardized to have unit variance. If not, the formulation becomes:

$$\sum_{i=1}^{k} \text{corr}^2(\mathbf{Z}\dashv, \mathbf{M}_i) = \frac{\dashv^\top \mathbf{Z}^\top \mathbf{M}\mathbf{D}^{-1}\mathbf{M}^\top \mathbf{Z}\dashv}{\dashv^\top \mathbf{Z}^\top \mathbf{Z}\dashv}, \tag{12}$$

where $\mathbf{D} = \text{diag}(\mathbf{M}_1^\top \mathbf{M}_1, \ldots, \mathbf{M}_k^\top \mathbf{M}_k)$.

This generalization allows for handcrafted weights for each KPI. In this case, $\mathbf{C}_1 = \mathbf{Z}^\top \mathbf{M}\mathbf{D}^{-1}\mathbf{M}^\top \mathbf{Z}$, and the argument still holds.

Now, $\mathbf{C}_1$ is symmetric by construction. For $\mathbf{C}_2 = \mathbf{Z}^\top \mathbf{Z}$ to be positive definite, $\mathbf{Z}$ must have full column rank. Since $\mathbf{Z}$ is the result of a singular value decomposition (SVD), this is guaranteed. Hence, the conditions of Theorem 1 are satisfied.

Therefore, the eigenvector $\dashv$ corresponding to the largest eigenvalue of the GEP $\mathbf{C}_1 \dashv = \lambda \mathbf{C}_2 \dashv$ is the optimal solution to Problem 6. □

IV. Results and Discussion

To determine whether our system performs better than the baseline (optimization with a single metric), the following A/B testing scenario is proposed:

1) Experiment Settings: We consider running a three-arm experiment comparing the following groups:

- **Group A (X-only)**: Ranker tuned only for KPI X.
- **Group B (Y-only)**: Ranker tuned only for KPI Y.
- **Group C (Multi-Metric)**: Ranker optimized jointly for X and Y.

Traffic should be randomly split into three equal buckets (approximately 33.3% each). For each group, the KPI metrics X and Y should be collected. It is recommended that the experiment runs for at least one week to ensure statistically meaningful and reliable results.

2) Evaluation Criteria: For each experiment group g, we define the percentage lift in KPI X as:

$$\Delta_X^g = \frac{\bar{X}_g - \bar{X}_{\text{baseline}}}{\bar{X}_{\text{baseline}}} \tag{13}$$

It is expected that:

- Group A should show a positive Δ_X^A and a neutral or negative Δ_Y^A.
- Group B should show a positive Δ_Y^B and a neutral or negative Δ_X^B.

Ideally, we expect to observe:

$$\Delta_X^C > \Delta_X^A \quad \text{and} \quad \Delta_Y^C > \Delta_Y^B$$

To **realistically** evaluate whether our system performs better than the current system, we define the *joint win-rate* metric as follows:

$$\delta_x = \frac{X_C - X_A}{X_A}, \quad \delta_y = \frac{Y_C - Y_B}{Y_B} \tag{14}$$

We then measure the fraction of sessions where both $\delta_x > 0$ and $\delta_y > 0$. This shows the fraction of the points where our proposed algorithm performs better than optimizing for a single metric.

V. Conclusion

...

References

[1] P. Shah, T. Chevalier *et al.*, "Online ranking with constraints: A primal-dual algorithm and applications to web traffic-shaping," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017.

[2] J. Yan, S. Chatterjee *et al.*, "Ads allocation in feed via constrained optimization," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2020.

[3] W. Ding, "Whole page optimization with global constraints," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019.

[4] L. Sun, X. Li *et al.*, "Traffic shaping in e-commercial search engine: Multi-objective online welfare maximization," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021.

[5] D. Carmel, L. Lewin-Eytan *et al.*, "Multi-objective ranking optimization for product search using stochastic label aggregation," in *Proceedings of The Web Conference 2020*, 2020.

[6] J. Tang, S. Katariya *et al.*, "Multi-objective learning to rank by model distillation," in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2024.

[7] D. Mahapatra, M. Momma *et al.*, "Multi-label learning to rank through multi-objective optimization," in *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2022.

[8] C. Shen, Q. Da *et al.*, "A general traffic shaping protocol in e-commerce," *arXiv preprint arXiv:2108.02805*, 2021.

[9] E. Malthouse, M. Fudurić *et al.*, "An algorithm for allocating sponsored recommendations and content: Unifying programmatic advertising and recommender systems," *Journal of Advertising*, 2019.

[10] I. Benouaret, J. Chagraoui *et al.*, "A bi-objective approach for product recommendations," in *2019 IEEE International Conference on Big Data (Big Data)*, 2019.

[11] Y. Zhu, Z. Chen *et al.*, "Optimizing search engine revenue in sponsored search," in *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, 2009.

[12] B. Roberts, P. Key *et al.*, "Ranking and tradeoffs in sponsored search auctions," in *ACM Conference on Economics and Computation*, 2013.

[13] L. Y. Chu, H. Zhang *et al.*, "Position ranking and auctions for online marketplaces," *Microeconomics: Production*, 2017.

[14] F. Radlinski, L. Riedel *et al.*, "Optimizing relevance and revenue in ad search: a query substitution approach," in *Proceedings of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2008.

[15] F. Pourbabaee, D. Mo *et al.*, "Trading off relevance and revenue in the jobs marketplace: Estimation, optimization and auction design," *arXiv preprint arXiv:2501.00000*, 2025.

[16] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.

[17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.