# NDBI021, Lecture 9

User preferences, 2/1 ZK+Z, Wed 12:20 - 13:50 S8 Wed 14:00 - 15:30 SW2 (odd weeks) https://www.ksi.mff.cuni.cz/~peska/vyuka/ndbi021/2022/



https://ksi.mff.cuni.cz

# Multi-criteriality in RS



https://ksi.mff.cuni.cz

# **Tutorial:**

## **Multi-Objective Recommendations**

Yong Zheng, Illinois Institute of Technology, USA David (Xuejun) Wang, Morningstar, Inc., USA





# **Different Types of Recommender Systems**



# Why MOO in RecSys

- There is an emerging demand in MOO
  - Traditional RecSys
    - Example: RecSys balancing multiple metrics, e.g., news



# Why MOO in RecSys

- There is an emerging demand in MOO
  - New Types of RecSys
    - Example: Multi-stakeholder RecSys, e.g., marketplace



# **RecSys with MOO**

• Contexts in which we need MOO in RecSys



#### Contents

- Background and Some History
- Multi Objective Optimization (MOO)
- MOO Solutions
  - Scalarization Algorithms
  - Multi Objective Evolutionary Algorithms
- Selection of the best solution in Pareto set
- MOO libraries
- Summary & QA



Multi Objective Optimization in Finance





- Multi Metrics in Recommendation Systems
  - Goal: Meets user's need
  - Objectives:
    - Maximize Accuracy
    - Maximize Diversity
    - Maximize Novelty
  - Challenge
    - Increase Diversity may decease Accurac
    - Increase Novelty may decease Accuracy





- Multi Stakeholder Recommendation Systems
  - Goal: Meet interests of all stakeholders
  - Objectives: Maximize Three Item Utilities
    - In respect of End User
    - In respect of Provider
    - In respect of Platform Owner
  - Challenge
    - Utilities regarding to three stakeholders may conflict each other





- Francis Ysidro Edgeworth (1845-1926)
  - Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences, published in 1881
  - "It is required to find a point (x, y) such that, in whatever direction we take an infinitely small step, P and  $\Pi$  do not increase together, but that, while one increases, the other decreases"
- Vilfredo Pareto (1848-1923)
  - -Manual of Political Economy, published in 1906

-"The optimum allocation of the resources of a society is not attained so long as it is possible to make at least one individual better off in his own estimation while keeping others as well off as before in their own estimation."







#### Contents

- Background
- Multi Objective Optimization (MOO)
- MOO Solutions
  - Scalarization Algorithms
  - Multi Objective Evolutionary Algorithms
- Selection of the best solution in Pareto set
- MOO libraries
- Summary & QA



#### Multi Objective Optimization (MOO) Objectives Multi Objective Optimization (MOO) Problem $\min(f_1(x), f_2(x), \dots, f_M(x))$ Subject to: $g_i(x) \ge 0, \quad j = 1, 2, ..., J$ Possible $h_k(x) = 0, \qquad k = 1, 2, \dots, K$ constraints format $x_i^L \le x_i \le x_i^U$ , i = 1, 2, ..., n**Decision variable:** $x \in \mathbb{R}^n$ **Objective Functions**: $f_i$ , i = 1, 2, ..., M**Feasible Solutions:** $S = \{x \mid x_i^L \le x_i \le x_i^U, g_i(x) \ge 0, h_k(x) = 0, j = 1, 2, ..., J, k = 1, 2, ..., K, i = 1, ..., n\}$ 13 M RNINGSTAR

Example: Two Objectives in Recommender Systems

Name	Symbol	Meaning
Decision Variable	X	Top N recommendation list
Feasible Solution Set	S	All top N recommendation list
First Objective	$f_1(x)$	1 — accuracy
Second Objective	$f_2(x)$	1 – diversity

Find recommendation list that maximize accuracy and diversity

$$\min_{x \in S} (f_1, f_2)$$
 or  $\min_{x \in S} F(x)$ , where  $F(x) = (f_1, f_2)$ 



- Special Characters of MOO
  - Objectives may be conflict each other
  - Cannot determine which solution is better
  - Example:

$$\min_{x}(f_{1}, f_{2})$$
  
Where  $f_{1}(x) = 2(x - 1) + 1$ ,  
 $f_{2}(x) = 2(x - 3)^{2} + 1$   
Subject  $x \in [0, 6]$ 





#### Dominance Relation

A solution x is said to be Dominated by  $x^*$  if and only if

$$f_m(\chi^*) \le f_m(\chi)$$
 for all  $m = 1, 2, ..., M$ 

and there exists at least one m' such that:

$$f_{m'}(x^*) < f_{m'}(x)$$

A and B dominate C, D is only dominated by A. A and B: no dominance relationship D and B: no dominance relationship





- Non-Dominated Solution (Pareto Optimal Solution)
  - Not dominated by any other solutions
  - Solution A, B, E and G are Pareto Optimal
- Pareto Optimal Set:
  - All x such that F(x) is on curve from F(E) to F(G)
- Pareto Front:
  - All F(x) on curve from F(E) to F(G)

 $\min_{x}(f_{1}, f_{2})$   $f_{2} = 1 - diversity$  F(E) - F(D) - F(C) F(A) - F(B) - F(G)  $f_{1} = 1 - accuracy$ 



Example:

$$\min_{x}(f_{1}, f_{2})$$
Where  $f_{1}(x) = 2(x - 1) + 1$ ,  
 $f_{2}(x) = 2(x - 3)^{2} + 1$ 
Subject  $x \in [0, 6]$ 

- Analysis
  - Feasible solutions: S = [0,6]
  - Pareto Set:  $\{x \mid x \in [0,3]\}$
  - Pareto Front:  $\{(f_1, f_2) | x \in [0, 3]\}$



Solving Multi Objective Optimization (MOO) Problem

 $\min F(\mathbf{x})$ 

where 
$$F(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_M(\mathbf{x}))$$
  
 $\mathbf{x} \in S$   
 $S$  is set of all feasible solutions

- Outputs
  - Find a Non-Dominated Solution
  - Find All Non-Dominated Solutions (Pareto Set)
  - Find a representative subset of Non-Dominated Solutions



MOO Decision Making Process





## Scalarization Algorithms

- Transform multi-objectives into a single objective
- Solve it by single objective optimizer
- Find one Pareto optimal solution in one run
- Find Pareto Set in multiple run
- Multi Objective Evolutionary Algorithms (MOEA)
  - Follow natural evolution process such as gene evolution, a flock of birds seeking food and other resources, a cooling process of melted crystal, ...
  - Find multiple Pareto optimal solutions in one run



#### Contents

- Background and Some History
- Multi Objective Optimization (MOO)
- MOO Solutions
  - Scalarization Algorithms
  - Multi Objective Evolutionary Algorithms
- Selection of the best solution in Pareto set
- MOO libraries
- Summary & QA



#### **Scalarization Algorithms**

- Weighting Methods
- ε-Constraint Method
- Normal Boundary Intersection (NBI) & Normal Constraint (NC)
- Goal Programming
- Physical Programming
- Lexicographic Method



#### Scalarization Algorithms: Weighting Methods

- Weighted Sum Method
  - A weight vector based on DM preference of each objectives:

$$\min_{x} \sum_{i=1}^{M} w_i f_i(x)$$
  
subject to  $x \in S$   
Where  $\sum_{i=1}^{M} w_i = 1$  and  $w_i > 1$ 

0

The condition of the weights guarantees Pareto optimal



Scalarization Algorithms: Weighted Sum Methods

Example: Two Objective Metrics Recommender Systems

 $\min_{x \in S} (f_1, f_2), \quad f_1 = 1 - \text{accuracy}, f_2 = 1 - \text{diversity}$ 

Solve:  $\min_{x \in S} (w_1 f_1 + w_2 f_2)$  $w_1 + w_2 = 1, w, w_2 > 0$ 



- Each  $(w_1, w_2)$  gives one Pareto solution
- Question: Can we get all Pareto solutions in this way? Spoiler: NO



#### **Scalarization Algorithms:** Weighting Methods Summary

#### Conditions of Pareto optimal solution

Method	Formula	Pro	Con
Weighted Sum	$\sum_{i=1}^{M} \mathbf{w}_i f_i(x)$	Simple	Require convex condition for Pareto set
Weighted Exponential Sums	$\sum_{i=1}^{M} \mathbf{w}_i [f_i(x)]^p$	Increase $p$ to approximate Pareto set	Bigger <i>p</i> may give non-Pareto solution
Weighted Metric Methods	$\left[\sum_{i=1}^{M} \mathbf{w}_{i}^{p}   f_{i}(x) - f_{i}^{*}  ^{p}\right]^{\frac{1}{p}}$	Different choice of ideal points	Bigger <i>p</i> may give non-Pareto solution
Weighted Chebyshev method	$\max_{i} \{ \boldsymbol{w_i}   f_i(x) - f_i^*   \}$	Can find complete Pareto set	Some solution may not Pareto optimal
Exponential Weighted Criterion	$\sum_{i=1}^{K} (e^{p\mathbf{w}_i} - 1)e^{pf_i(x)}$	Can find complete Pareto set	May lead computation over low
Weighted Product Method	$\prod_{i=1}^{K}  f_i(x) ^{w_i}$	Deal with different magnitude of objectives	Rarely used

## **Scalarization Algorithms**

*ϵ*-Constraint Method<sup>1</sup>

 $\min f_{l}(x)$ subject to  $f_{i}(x) \leq \epsilon_{i}$ , for all  $i \neq l$  $\epsilon_{i}$  is a known the upper bound of  $f_{i}$ 

- 1) Choose different  $\epsilon_i$  may produce all Pareto solutions
- 2) No convex requirement

#### 3) May not Pareto optimal

1. Haimes, Lasdon, Wismer, On a Bicriterion Formulation of the Problems of Integrated, System Identification and System Optimization, IEEE Transactions on Systems, Man, And Cybernetics, July 1971



#### **Scalarization Algorithms:** $\epsilon$ -Constraint Method

- A sufficient condition of Pareto optimal<sup>1</sup>
  - If optimal solution  $x^*$  is unique
- Two objective example

#### solve

 $\min_{x \in S} f_1(x)$ <br/>subject to  $f_2 \leq \epsilon$ 

Constraint	Unique Solution	Pareto Optimal
$\epsilon^1$	No solution	NA
$\epsilon^2, \epsilon^3, \epsilon^4, \epsilon^5$	Yes	Yes
$\epsilon^6$	No	Not necessary



1. V. Chankong, Y. Haimes, Multiobjective Decision Making, Dover Publication, 1983



 $\epsilon^6$ 

#### Contents

- Background and Some History
- Multi Objective Optimization (MOO)
- MOO Solutions
  - Scalarization Algorithms
  - Multi Objective Evolutionary Algorithms
- Selection of the best solution in Pareto set
- MOO libraries
- Summary & QA



#### **Multi Objective Evolutionary Algorithms**

Evolutionary Algorithms inspired by natural evolutionary process:





#### **Multi Objective Evolutionary Algorithms**

Benefits:



Objective can be any function



Parallel computing





#### Multi Objective Evolutionary Algorithms : Basic Concepts

- Terminologies of Solutions:
  - Individual : a feasible solution x
  - Population: a set of individuals
  - Parents: selected from Population
  - Children: produced from Parents





#### Multi Objective Evolutionary Algorithms : Basic Concepts

- Operators of Genetic Algorithm
  - Evaluation: measure how 'good' each solution is
    - assigning fitness value (or order): z(x)
  - Selection: find Parents
    - Random process
    - Tournament process
  - Variation: produce children
    - Crossover
    - Mutation
  - Elitism:
    - maintain 'better' solution in each iteration



#### Multi Objective Evolutionary Algorithms : Encoding in Genetic Algorithm

- Binary Encoding in Recommender
  - Chromosome: a recommendation list (decision variable x)
  - Length of Chromosome = total number of available items
  - Each Gene position is corresponding an item
    - 1: item is in recommender list,
    - 0: item is not in recommender list





#### Multi Objective Evolutionary Algorithms : Encoding in Genetic Algorithm

- Permutation Encoding in Recommender
  - Length of Chromosome = N in top N recommendation list
  - The value of each Gene: index of an item
    - 6: the 6<sup>th</sup> item
    - Total 9 available items

-N = 6





## **From Single Objective to Multi Objective Evolutionary Algorithms**

- Solve a MOO problem
  - Find non-dominated solutions (Pareto optimal)
- Where are dominance relations applied? Where are multi objective considered?
  - Fitness value z(x) evaluation
  - Parent selection
  - Elitism





#### **Multi Objective Evolutionary Algorithms:**

Major MOO Genetic Algorithm Methods

Method	Fitness Evaluation	Parent Selection	Elitism
VEGA (Schaffer,1985)	Single objective	Probability distribution	Dominance relation
MOGA (Fonseca & Fleming, 1993)	Dominance relation	Probability distribution	NA
NSGA (Srinivas and Deb,1994)	Dominance relation	Probability distribution	NA
NSGA-II (Debb, etc., 2002)	Dominance relation	Probability distribution	<b>Dominance</b> relation
NPGA (Horn, etc. ,1994)	No Fitness	Tournament method (dominance relation)	NA
PAES (Knowles and Corne, 1999)	No Fitness	Local search (dominance relation)	<b>Dominance</b> relation

## **Classification of Multi Objective Genetic Algorithms**

- Dominance relation in Elitism
  - VEGA (Schaffer, 1985)
- Dominance relation in fitness function
  - MOGA (Fonseca & Fleming, 1993),
  - NSGA (Srinivas and Deb,1994),
  - NSGA-II (Debb, etc., 2002)
- No fitness values, but dominance relation in selection process
  - NPGA (Horn, etc., 1994),
  - PAES (Knowles and Corne, 1999)



## **Multi Objective Evolutionary Algorithms**

56



1. Schaffer, 1985: *Multiple Objective Optimization with Vector Evaluated Genetic Algorithms*, The First International Conference on Genetic Algorithms and their Applications (held in Pittsburgh), pp. 93–100



## **Classification of Multi Objective Genetic Algorithms**

- Fitness value determined by single objective
  - VEGA (Schaffer, 1985)
- Fitness value determined by dominance relations
  - MOGA (Fonseca & Fleming, 1993),
  - NSGA (Srinivas and Deb,1994),
  - NSGA-II (Debb, etc., 2002)
- No fitness value needed:
  - NPGA (Horn, etc., 1994),
  - PAES (Knowles and Corne, 1999)



## **Multi Objective Evolutionary Algorithms**

- Nondominated Sorting Genetic Algorithm (NSGA<sup>1</sup>)
  - Sorting by selecting non-dominated solutions without replacement each time:



1. Srinivas and Deb, 1994: Multiobjective optimization using nondominated sorting in genetic algorithms. Journal of Evolutionary Computing



#### **Multi Objective Evolutionary Algorithms**

#### MOEA methods based on other EAs

#### Particle Swam Optimization (PSO)

James Kennedy and Russell C. Eberhart. *Particle swarm optimization*. Proceedings of the 1995 IEEE International Conference on Neural Networks, Piscataway, New Jersey, 1995

M. Reyes-Sierra and C. Coello, *Multi-Objective Particle Swarm Optimizers: A Survey of the State-of-the-Art*, International Journal of Computational Intelligence Research, Vol.2, No.3 (2006), pp. 287–308



#### Contents

- Background and Some History
- Multi Objective Optimization (MOO)
- MOO Solutions
  - Scalarization Algorithms
  - Multi Objective Evolutionary Algorithms
- Selection of the best solution in Pareto set
- MOO libraries
- Summary & QA



- When do we need to produce Pareto set?
  - Most MOEA method (A posterior, No DM available)
  - Scalarization without DM preference (A posterior, No DM available)
- Example: Recommender Systems balancing multi metrics
  - End user (DM) cannot choose from Pareto set
  - A single best recommendation list needs to be produced



- No information from decision maker (DM)
  - All solutions in Pareto set are 'equally good'
  - Need to make the 'best' guess
- Best guess method
  - Knee point method
  - Hypervolume Method
  - Multiple-criteria decision-making (MCDM) methods

 $f_2$   $A_1$  $A_2$  $A_3$  Which one is better?  $A_4$  $A_4$  $A_5$   $A_6$   $A_7$   $A_8$ 



- Knee Point
  - A special point (A) on Pareto Front
  - small improvement in either objective will cause a large deterioration in the other objective,  $f_2^{\dagger}$
- Find Knee Point
  - Angle Based Method<sup>1</sup>
  - Marginal Utility Method<sup>1</sup>
  - Hyperplane Normal Vector Method<sup>2</sup>



1. Deb and Gupta, Understanding Knee Points in Bicriteria Problems and Their Implications as Preferred Solution Principles, Engineering Optimization, 43(11) 2. Yu, Jin, Olhofer, A Method for a Posteriori Identification of Knee Points Based on Solution Density, 2018 IEEE Congress on Evolutionary Computation (CEC)



- Angle Based Knee Point<sup>1</sup>
  - Only work in two objectives MOO
  - Pareto front:  $\{A_1, A_2, \dots, A_8\}$
  - Calculate reflex angle  $\alpha$  for each point:

 $\alpha_i = \angle A_{i-1}A_iA_{i+1}$ 

– Find the point with  $\max \alpha_i$ 

 $\alpha_5 = \max\{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7\}$ 

#### $A_5$ is the Knee point



1. Deb and Gupta, Understanding Knee Points in Bicriteria Problems and Their Implications as Preferred Solution Principles, Engineering 70 Optimization, 43(11)



# **RecSys with MOO**

• Contexts in which we need MOO in RecSys



- Why we need MOO in this context
  - Relevance or accuracy is not the only focus
    - For example, news and music recommendations
      - Boring if always recommending the same types of items
      - Diversity: try something different
      - Novelty: try something never experienced before
    - For example, item recommendations in e-commerce
      - Co-sales
      - Bundle sales

- Goals
  - Improve other metrics at no loss or acceptable loss on accuracy
  - Challenges
    - No clear rules to define the "acceptable" loss

• Case Study 1: Hybrid Recommender

Ribeiro, M. T., Lacerda, A., Veloso, A., & Ziviani, N. (2012). Pareto-efficient hybridization for multi-objective recommender systems. In ACM RecSys 2012.

- Application: balancing accuracy, novelty, diversity
- Recommendation Framework



## • Case Study 1: Hybrid Recommender

Ribeiro, M. T., Lacerda, A., Veloso, A., & Ziviani, N. (2012). Pareto-efficient hybridization for multi-objective recommender systems. In ACM RecSys 2012.

## – MOEA as the MOO Method

- Consider accuracy, diversity, novelty as objectives
- Use Strength Pareto Evolutionary Algorithm as MOEA optimizer
  - Encoding/Decision variables: the weights in the hybrid model
  - Output: a Pareto optimal set
- Select the best single solution from Pareto set
  - Use a weighted sum on the three objectives
  - Try different set of weights  $(Q_j)$  manually
- Results: balancing multiple metrics



# **RecSys with MOO**

• Contexts in which we need MOO in RecSys



- Why we need MOO in this context
  - The end user is not the only stakeholder



 RecSys should be built by considering the item utility from the perspective of different stakeholders

- Objective definitions
  - It varies from domains to domains
  - For each stakeholder, there's at least one objective
    - E-Commerce or Marketplace
      - Buyer: user preferences on items, budget
      - Seller: profits
      - Platform: commission fees
      - Delivery company: costs and profits
    - Job seeking
      - Job seeker: user preferences
      - Recruiter: talent requirements

## • Goals

- Deliver item recommendations by balancing the needs of multiple stakeholders
- With acceptable loss on the consumer side
- Challenges
  - Which stakeholders should be considered
  - How to define and achieve the "balance"
  - No clear rules to define the acceptable loss

• Case Studies

#### – Using scalarization as the MOO method

Lin, X., Chen, H., Pei, C., et al. (2019). A pareto-efficient algorithm for multiple objective optimization in e-commerce recommendation. In ACM RecSys, 2019.

### – Using MOEA as the MOO method

Zheng, Y., Ghane, N., & Sabouri, M. (2019). Personalized educational learning with multi-stakeholder optimizations. In Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization (pp. 283-289).

- Case Study 1: Using scalarization in E-Commerce
  - Objectives
    - CTR (Click Through Rate)

Lin, X., Chen, H., Pei, C., et al. (2019). A pareto-efficient algorithm for multiple objective optimization in e-commerce recommendation. In ACM RecSys, 2019.

GMV (Gross Merchandise Volume)



- Case Study 1: Using scalarization in E-Commerce
  - MOO Method
    - Define a loss function for each objective

CTR  $\mathcal{L}_{CTR}(\theta, \mathbf{x}, y, z) = -\frac{1}{N} \sum_{j=1}^{N} log(P(y_j | \theta, \mathbf{x}_j))$ , i.e., point-wise learning-to-rank GMV  $\mathcal{L}_{GMV}(\theta, \mathbf{x}, y, z) = -\frac{1}{N} \sum_{j=1}^{N} h(price_j) \cdot log(P(z_j = 1 | \theta, \mathbf{x}_j))$ 

x: impression, y: clicks, z: purchases

• Use weighted sum as the scalarization Joint Loss =  $\omega \cdot L_{CTR} + (1 - \omega) \cdot L_{GMV}$ 

- Case Study 1: Using scalarization in E-Commerce
  - MOO Method
    - Use weighted sum as the scalarization Joint Loss =  $\omega \cdot L_{CTR} + (1 \omega) \cdot L_{GMV}$
    - Try different weights to get Pareto Set
    - Select a single best solution by using Least Misery strategy, i.e., minimizing the highest loss function of the objectives min max{L<sub>1</sub>, L<sub>2</sub>,..., L<sub>K</sub>}

## Summary

Suggested Workflow



So, who is the decision maker in RecSys?

- The end user
- The developer



## Summary

- Our Tutorial
  - Website: <u>https://moorecsys.github.io/</u>
  - Slide: <u>https://github.com/moorecsys/moorecsys.github.io</u>
  - Supplementary materials:

"Multi-Objective Recommendations: A Tutorial" on arXiv.org (will be available soon on the github above)