بسم الله الرحمن الرحبم

Optimization The Mathematical Language for Decision Making in Today's Complex World

Introduction to Optimization

History

- What is Optimization?
- Real-world relevance: logistics, finance, engineering, etc.

Classifications of optimization problems

- Linear programming
- Integer programing
- Non-linear programming
- Stochastic programming

Application

portfolio optimization

My research area



operations research is the application of methods of science to complex problems arising in the direction and management of large number of men machines materials and money in industry business government and defense also operations research is the application of scientific methods techniques and tools to the problems involving the operations of a system so as to provide those in control of the system with optimum solution to the problem





- complex logistical problems
- invention of a new flight patterns
- planning sea mining
- effective utilization of electronic equipment



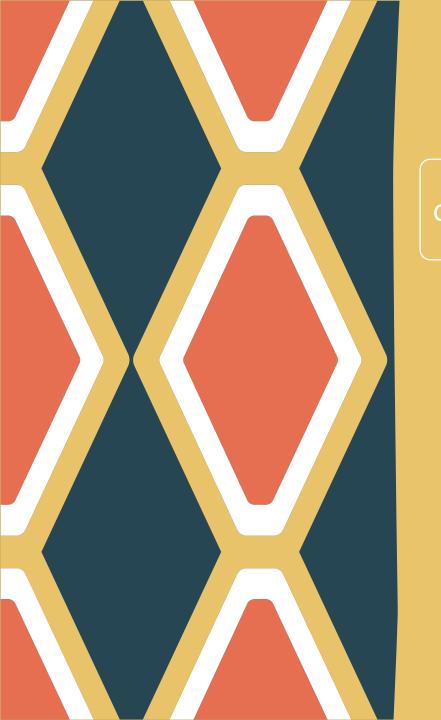


with the success of the military operations the industry then followed in using operations research this is to sought out solutions to their complex executive type problems to serve the overall objectives of their organizations and utilization of the effective tools

Military
 Business applications,
 hospitals,
 financial institutions,
 libraries,
 city planning,
 transportation systems,
 crime investigations.

 Sought solutions to their complex executive-type problems to serve the overall objectives of the organization utilization of the effective tools





Phases of study

definition of the problem

construction of the model

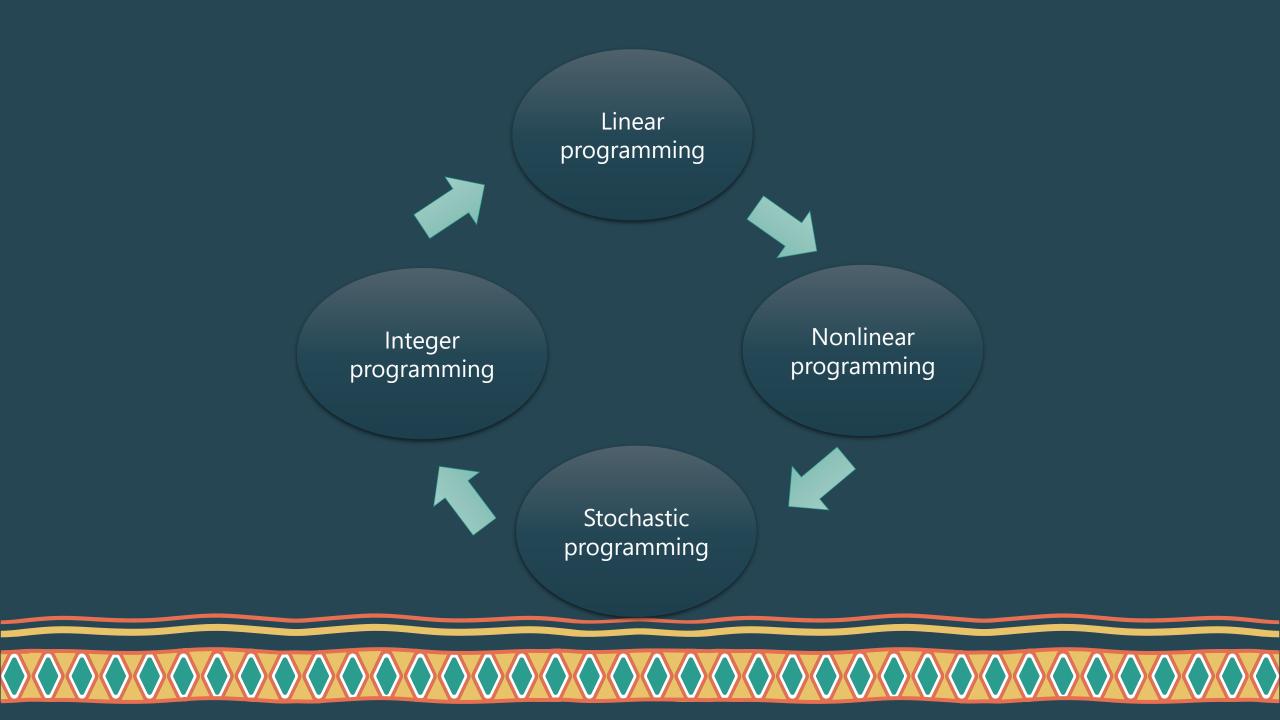
solution of the model

validation of the model

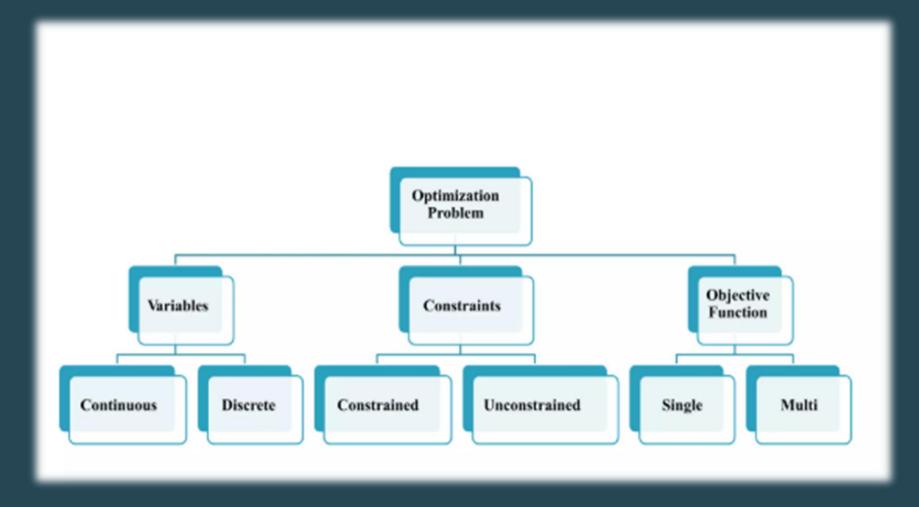
implementation of the final results

Classification of Optimization Models





Components of an optimization problems



What is linear programming or Lp

The maximization or minimization of some quantity is the objective in all linear programming problems. All LP problems have constraints that limit the degree to which the objective can be pursued. A feasible solution satisfies all the problem's constraints. An optimal solution is a feasible solution that results in the largest possible objective function value when maximizing (or smallest when minimizing). A graphical solution method can be used to solve a linear program with two variables.

If both the objective function and the constraints are linear, the problem is referred to as a linear programming problem. Linear functions are functions in which each variable appears in a separate term raised to the first power and is multiplied by a constant (which could be 0). Linear constraints are linear functions that are restricted to be "less than or equal to", "equal to", or "greater than or equal to" a constant.

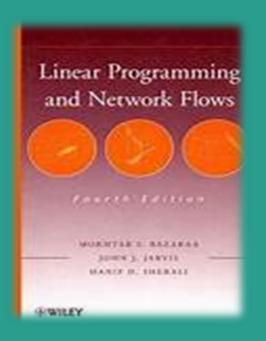
max \min

$$z = c_1 x_1 + \dots + c_n x_n$$

subject to:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n (\leq = \geq) b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n (\leq = \geq) b_2 \\ & \cdot \\ & \cdot \\ a_{m1}x_1 + \dots + a_{mn}x_n (\leq = \geq) b_m \end{cases}$$

 x_j = decision variables b_i = constraint levels c_j = objective function coefficients a_{ij} = constraint coefficients





areas of application of linear programming

industrial application

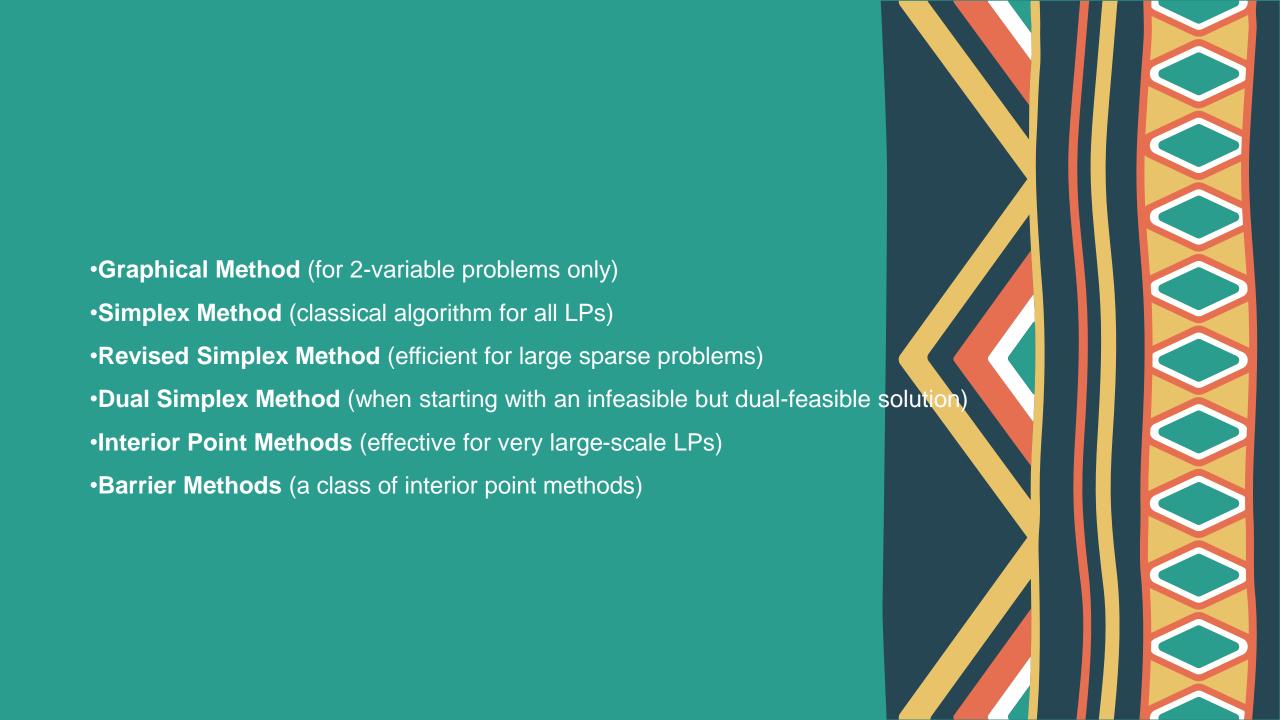
- products mix problem
- blending problems
- production scheduling problem
- · assembly line balancing
- make-or-buy problems

management applications

- media selection problems
- portfolio selection problems
- profit planning problems
- transportation problems

miscellaneous applications

- diet problems
- agriculture problems
- flight scheduling problems
- facilities location problems



limitation of linear program model

linear programming is applicable only to the problem where the constraints and objective function are linear i.e., where they can be expressed as an equations which represent straight lines. in area life situation, when constraints or objective function are not linear, this technique cannot be used

- factors such as uncertainty, and time are not taken into consideration.
- parameters in the model are assumed to be constant but in real life situation they are not constants.
- linear program deals with only single objective, whereas in the real-life situation may have multiple and conflicting objectives.
- in solving a Lp there is no guarantee that we get an integer value. in some case of no of men/machine and non-integer value is meaningless



Integer Programming (IP / ILP / MILP)

An integer programming model is one where one or more of the decision variables has to take on an integer value in the final solution. Solving an integer programming problem is much more difficult than solving an LP problem. Even the fastest computers can take an excessively long time to solve big integer programming problems. If requiring integer values is the only way in which a problem deviates from a linear programming formulation, then it is an integer programming (IP) problem. (The more complete name is integer linear programming, but the adjective linear normally is dropped except when this problem is contrasted with the more esoteric integer nonlinear programming problem. So, The mathematical model for integer programming is the linear programming model with the one additional restriction that the variables must have integer values..



Maximize $\sum_{j=1}^{N} c_j$

subject to:

$$\sum_{j=1}^{n} a_{ij} x_j = b_i \quad (i = 1, ..., n)$$

$$x_j \ge 0$$
 $(j = 1, ..., n)$

 x_i integer (for some or all j = 1, 2, ..., n).

PROGRAMMING PROBLEMS

- PURE-INTEGER PROBLEMS Require that all decision variables have integer solutions.
- MIXED-INTEGER PROBLEMS Require some, but not all, of the decision variables to have integer values in the final solution, whereas others need not have integer values.
- 0–1 INTEGER PROBLEMS Require integer variables to have value of 0 or 1, such as situations in which decision variables are of the yes- no type.

Common Solution Methods

1. Branch and Bound

Systematically explore solution tree

<u>Bound non-integer relaxations to prune subtrees</u>

2. Cutting Plane Method

Add linear inequalities (cuts) to eliminate fractional solutions

3. Branch and Cut

Combines Branch & Bound with Cutting Planes (used in modern solvers)

4. Heuristics & Metaheuristics

Greedy, Genetic Algorithms, Simulated Annealing, Tabu Search

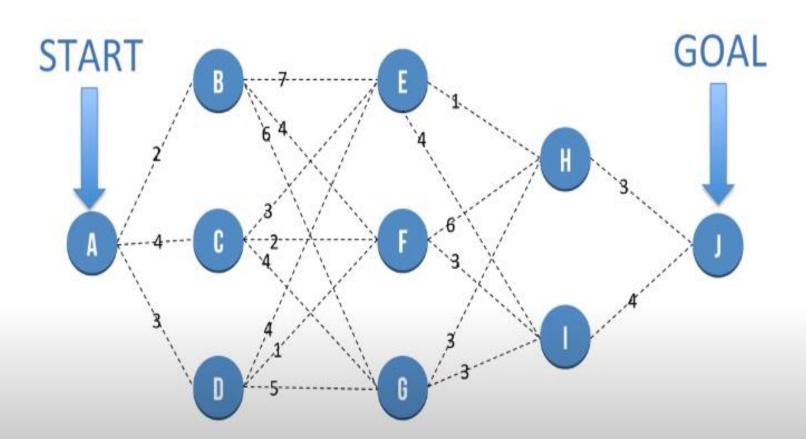
5. LP Relaxation

Relax integer constraints → solve as LP

6. Dynamic programing → solve substructure

Dynamic Programming Basic Idea – shortest path problem DP Solution

Principle of Optimality





In Finance

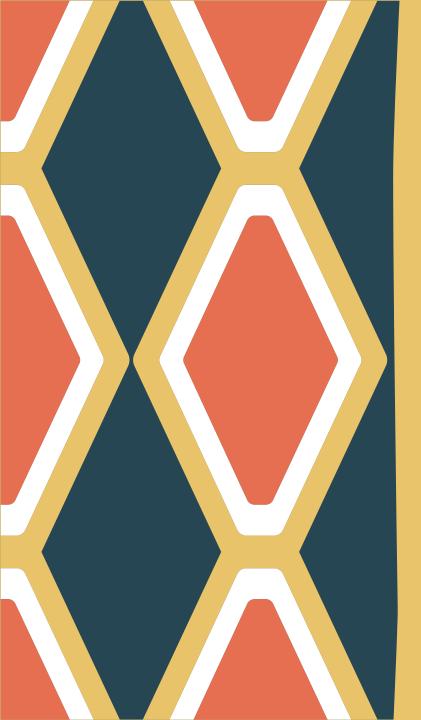
- Portfolio optimization with cardinality constraints
- Limit the number of assets in a portfolio
- Include transaction costs (piecewise linear, binary decisions)
- Option pricing with discrete trading rules

Operations

- Facility location problems: Where to open warehouses (binary variables)
- Production planning: Which products to produce, how much
- Scheduling: Assigning jobs to machines or workers
- •Routing: Vehicle routing problem (VRP), traveling salesman problem (TSP)

Supply Chain

- Supply allocation
- Inventory management
- Transportation with truck/batch limits



Nonlinear Programming (NLP)

A Nonlinear Programming (NLP) problem is an optimization problem where at least one of the following is nonlinear:

- The objective function
- One or more constraints

This makes NLP problems more complex than linear programming problems and often requires specialized solution techniques

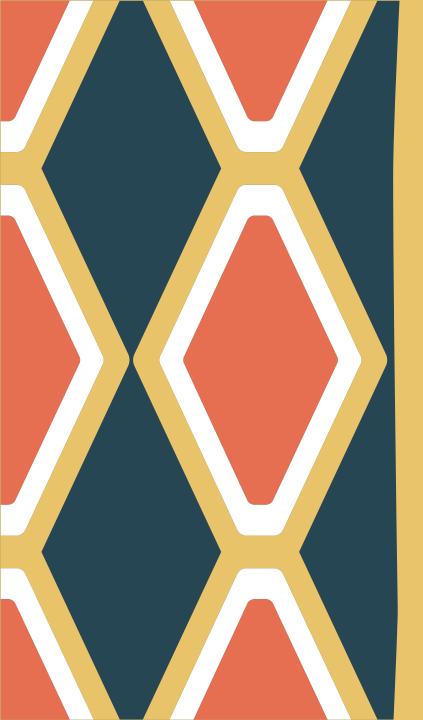
Why NLPs Are Harder Than LPs:

- •The feasible region may be non-convex \rightarrow can have local vs global optima.
- Solving requires iterative methods and initial guesses.
- •No guarantee of finding a global optimum without specific conditions (e.g., convexity).



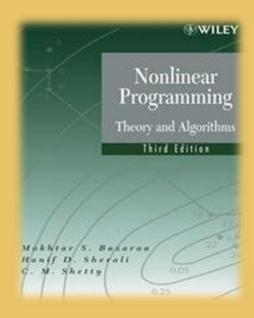
A general nonlinear programming problem (NLP) can be expressed as follows: Find the values of decision variables $x_1,...,x_n$ that

where $f(x_1,...,x_n)$ is the NLP's objective function, and $g_1(x_1,...,x_n)$ (\leq , =, or \geq) b_1 are the NLP's constraints. An NLP with no constraints is an unconstrained NLP. The feasible region for NLP above is the set of ($x_1,...,x_n$) that satisfy the m constraints in the NLP. A point in the feasible region is a feasible point, and a point that is not in the feasible region is an infeasible point.



Common NLP Solution Methods:

- **1.Gradient Descent / Steepest Descent**
- 2. Newton's Method / Quasi-Newton
- 3. Sequential Quadratic Programming (SQP)
- **4.Interior Point Methods**
- **5.Penalty and Barrier Methods**
- 6.Metaheuristics (e.g., Genetic Algorithms,
- 7. Simulated Annealing—for non-convex or non-differentiable NLPs)





Revenue and Pricing
Optimization

- portfolio selection problems (Markowitz Mean-Variance Model)
- make-or-buy problems
- nonlinear demand functions

Healthcare applications

- Radiation therapy optimization
- hospital resource allocation
- transportation problems

Machine Learning & Data Science

- Support Vector Machines with nonlinear kernels
- Neural network training (non-convex loss functions)
- Regularized logistic regression

Stochastic programming

Stochastic Programming is a mathematical optimization framework for decision-making under uncertainty. Unlike deterministic models, SP considers that some parameters (data) in the model are random variables with known probability distributions.

- Decision making under uncertainty
- Very general class of problems:
 - How to create and manage a portfolio
 - Optimal investment sequences, given
 - Historic distribution of returns and covariances
 - -Horizon, financial goals, regulatory constraints, etc.
 - -How to harvest a forest
 - Optimal harvest sequence, given
 - -Random incidence of forest fires, pest, etc.
 - -How to generate power
 - Random data on demand, rates, parameters
 - -etc.

Common characteristics

- Large-scale optimization models
- Some problem parameters unknown
- Assume distribution of parameters known
- (Otherwise: Optimization under risk)

Let x be the first-stage (here-and-now) decision variables, and $y(\xi)$ be the second-stage (recourse) decisions depending on the realization (ξ) of a random variable:

$$min_{x \in X} \quad \{c^T x + E_{\xi}[Q(x, \xi)]\}$$

where:

$$Q(x,\xi) = \min_{y} \{ q(\xi)^T y \mid T(\xi)x + W(\xi)y \ge h(\xi) \}$$

- Stage 1: Decide before uncertainty is revealed (e.g., how much to invest)
- Stage 2: Make adjustments after observing uncertainty (e.g., hedge losses)

Techniques

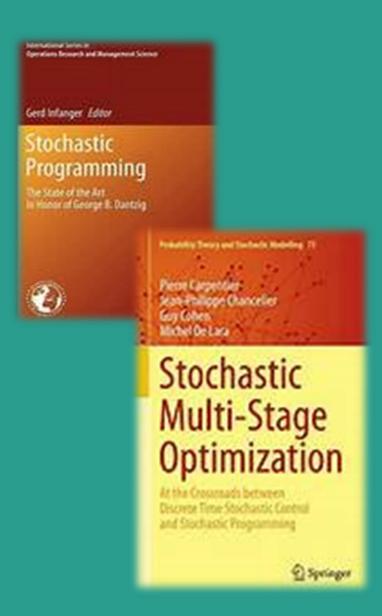
- 1. Scenario-Based Approach
 - Discretize uncertainty using a finite number of scenarios

 Solve a large deterministic equivalent program (DEP)

 Used in two-stage SP models
- 2. Chance constraint

$$P(g(x,\xi) \leq 0) \geq 1-\alpha$$

- -Individual Chance Constraints
- -Joint Chance Constraints
- 3. Robust optimization
- 4. Distributionally robust optimization



Energy Systems Planning

- · Unit commitment with uncertain demand
- Wind power integration with stochastic availability
- Stochastic OPF (Optimal Power Flow)

Supply Chain Management

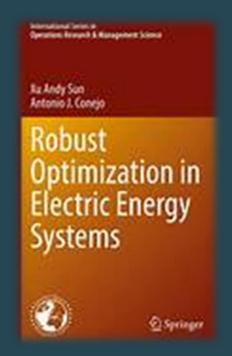
- Uncertain demand or lead time
- Optimal inventory and ordering policies
- Multi-stage production and distribution

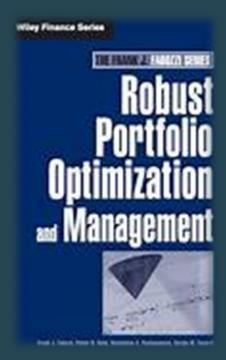
Others

- Machine Learning
- Project Scheduling
- Disaster Response & Risk Management

Robust Optimization

Robust Optimization is a mathematical framework for making decisions under uncertainty where the exact distribution of uncertain parameters is not known but bounded within an uncertainty set. The goal is to find solutions that perform well for all possible realizations of the uncertain parameters within these bounds.





Robust Optimization

It is a collection of methodologies for performing optimization when (some) of our data is uncertain and poorly understood. Suppose we have a nominal optimization problem:

$$min_{x \in \chi} f(x, u)$$

where u is an uncertain parameter. In Robust Optimization, instead of optimizing for a single value of u, we optimize for the worst-case scenario within a given uncertainty set u:

$$min_{x \in \chi} max_{u \in \mathcal{U}} \quad f(x, u)$$

Decision variables: $x \in \mathbb{R}^n$

Uncertainty set: $\mathcal{U} \in R^m$ (polyhedral, ellipsoidal, or general) Objective: Immunize against worst-case realizations within \mathcal{U}

Designing portfolios that perform well even in adverse market

Allocating assets under uncertain returns and covariance estimates

conditions

Robust counterparts of the Markowitz mean-variance model

Energy Systems

Others

Finance and **Portfolio**

Optimization

Route planning with uncertain travel times

Airline crew scheduling and vehicle routing problems

Designing transportation networks resilient to disruptions

Machine Learning

Healthcare and Medicine

Transportation and Logistics

Distributionally Robust Optimization (DRO)

Distributionally Robust Optimization (DRO) is a mathematical optimization framework designed to make decisions under uncertainty when the exact probability distribution of uncertain parameters is not fully known. Instead of assuming a single known distribution (as in stochastic programming), DRO considers that the true distribution lies within a set of plausible distributions, known as the ambiguity set.

$$min_{x \in X} sup_{P \in \mathcal{D}} \quad E_P[f(x, \xi)]$$

 $x \in X$: Decision variable in feasible set.

 ξ : Uncertain parameter.

 $f(x,\xi)$: Cost (or loss) function.

 \mathcal{D} : Ambiguity set containing all distributions considered plausible based on partial information (e.g., moments, Wasserstein distance, statistical samples).

 $sup_{P\in\mathcal{D}}$: Worst-case expected cost over all distributions in the ambiguity set.

Some Types of Ambiguity Sets

- •Moment-based sets: defined by mean, variance, and higher moments.
- •Wasserstein balls: distributions within a Wasserstein distance from an empirical distribution.
- φ-divergence sets: using KL-divergence, total variation, etc.

Balances robustness and performance. Offers distributional guarantees. Avoids overfitting to empirical samples. More realistic than assuming known distributions. Results in convex and tractable reformulations in many practical cases.

Applications areas

Finance and Portfolio Optimization:

- -create portfolios that remain effective across a range of market return distributions.
- -Models' portfolio returns using ambiguity sets to control worst-case risk (e.g., DRO mean-CVaR portfolio).

Energy Systems

- -Planning under uncertainty in demand, price, and renewable generation.
- -Ensures system feasibility even under extreme distributional deviations.

Transportation and Logistics

- -Designing routes and schedules with uncertain travel times and demand.
- -Models that incorporate data-driven ambiguity sets improve reliability.

Others

- -Robust facility location, staff planning, and manufacturing scheduling.
- -Enables robust design of systems facing uncertain cost or performance metric

Portfolio optimization

1. Markowitz Mean-Variance Theory (1952) Trade-off between risk and return.

$$min_x \ x^T \Sigma x$$
 $s.t: \ x^T \mu \ge \mu_0, \ \sum x_i = 1, \ x_i \ge 0$

- *x*: portfolio weights
- μ : expected return vector
- Σ : covariance matrix of returns
- •Assumptions:
 - Returns are normally distributed
 - Risk measured by variance
 - Investors are risk-averse and rational

This formed the basis of Modern Portfolio Theory (MPT) and introduced the efficient frontier.

- 2. Capital Asset Pricing Model
 - -Arbitrage Pricing Theory
- 3. Extensions and Constraints
 - **-Quadratic Programming with Constraints**
 - .Cardinality constraints (e.g., max number of assets)
 - .Minimum/maximum holding bounds
 - .Turnover constraints (transaction cost control)
 - -Robust Optimization
 - .Began incorporating uncertainty in inputs (mean, covariance).

4. Revolutions and Modern Approaches

Robust Portfolio Optimization

- -Deals with parameter uncertainty, especially in μ and Σ
- -Uses uncertainty sets and worst-case optimization:

 $\min_{\mathbf{x}} max_{\mu \in \mathcal{U}, \Sigma \in \mathcal{U}} \mathbf{x}^{\mathrm{T}} \Sigma \mathbf{x} - \lambda \mathbf{x}^{\mathrm{T}} \mu$

Risk-Based Approaches

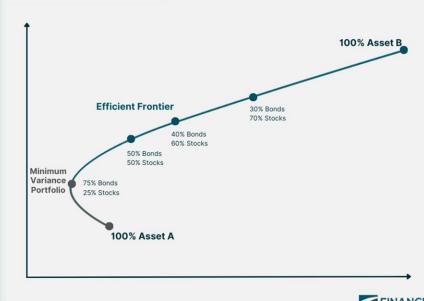
- -Avoid expected return due to its instability.
- -Emphasize risk budgeting:
- -Minimum-Variance Portfolio
- -Equal Risk Contribution (ERC)
- -Maximum Diversification

Scenario-Based and Stochastic Programming

-Use historical or simulated scenarios of returns.

 $\min_{\mathbf{x}} \mathbf{E}_{\boldsymbol{\omega}} \mathbf{L}(\mathbf{x}, \boldsymbol{\omega})$

The **Efficient Frontier**





Machine Learning in Portfolio Selection Learning-based methods for return forecasting and asset selection:

- Reinforcement Learning
- Deep Learning for dynamic allocation
- Clustering and dimensionality reduction

Distributionally Robust Optimization (DRO)

- -Optimizes over worst-case distributions within an ambiguity set.
 - -Wasserstein DRO
 - -Moment-based DRO

Reinforcement Learning and Online Optimization

- -Adapts to market changes dynamically.
- -Learns portfolio policies over time with minimal supervision.

Application in portfolio optimization

Challenge: In real-world portfolio optimization, the true probability distribution of asset returns is unknown. Traditional stochastic optimization assumes full knowledge, leading to fragile solutions.

Goal: Develop a robust and flexible model that accounts for distributional ambiguity and ensures superior out-of-sample performance.

Approach: Use a distributionally robust optimization (DRO) framework under a Wasserstein ambiguity set, incorporating second-order stochastic dominance (SSD) constraints to ensure risk-averse preferences are respected.

Stochastic dominance

Stochastic dominance is a partial order between random variables.[1][2] It is a form of stochastic ordering. The concept arises in decision theory and decision analysis in situations where one gamble (a probability distribution over possible outcomes, also known as prospects) can be ranked as superior to another gamble for a broad class of decision-makers. It is based on shared preferences regarding sets of possible outcomes and their associated probabilities. Only limited knowledge of preferences is required for determining dominance. Risk aversion is a factor only in second order stochastic dominance.

$$P(A \ge x) \ge P(B \ge x)$$
 (FSD)

$$E[g(y)] \le E[g(x)] \tag{SSD}$$

$$m = \{ \mu \in P(\Omega) : w^r(\mu, \nu) \le \varepsilon, E_{\mu}[(\xi - m_0)(\xi - m_0)^T] \le \Sigma_0 \}$$

$$max_x \quad E[\xi^T x]$$
 $s.t. \quad \sum_{i=1}^n x_i = 1$
 $E(\gamma - [\xi^T x]) \le E(\gamma - Y)$ **step 1**
 $x_i \ge 0$

$$\sup_{x \in R^n inf_{\mu \in m}} E[\xi^T x]$$

 $s.t. \sup_{\xi} E(\gamma - [\xi^T x])_+ \le E(\gamma - Y)_+$
 $\sum_{i=1}^n x_i = 1$
 $x > 0$

Methodological Contributions

Ambiguity Set: Defined using the Wasserstein distance, it includes all distributions close to the empirical one, providing high confidence via the concentration of measure theorem.

Objective: Maximize worst-case expected return over all distributions in the ambiguity set.

Constraints: Enforce ambiguous SSD dominance—ensuring the portfolio return stochastically dominates a benchmark.

Reformulations:

- Semidefinite Programming (SDP)
- Second-Order Cone Programming (SOCP)
- Semi-infinite programming solved via cutting surface and cutting plane methods.

Validation: Extensive experiments confirm robustness and efficiency.

Real-World Impact

Application: Portfolio optimization with 30-industry real stock data. Benefits:

- Accounts for distributional uncertainty without overfitting.
- Ensures risk-averse investment behavior using SSD.
- Outperforms classical and moment-based models in return consistency.
- •Innovation: Combines distributional robustness with stochastic dominance and the Wasserstein metric, enabling better risk-return trade-off and data-driven decision making in finance.
- •Real-World Problem Solved: Provides a principled and scalable solution to robust investment planning under deep uncertainty, relevant to institutional investors, pension funds, and financial risk managers.

My Research Focus

- Topic: Portfolio Optimization under Uncertainty
- •Techniques:
 - Second-order Cone Programming (SOCP)
 - Distributionally Robust Optimization
 - Wasserstein Distance for Ambiguity Sets
- •Objectives:
 - Minimize CVaR or worst-case loss
 - Balance return and risk with real-world data
- •Application:
 - Stock markets, pension fund management, algorithmic trading



Information Sciences



journal homepage: www.elsevier.com/locate/ins



Distributionally robust portfolio optimization with secondorder stochastic dominance based on wasserstein metric



Zohreh Hosseini-Nodeh 4.*, Rashed Khanjani-Shiraz 4.*, Panos M. Pardalos b

*Faculty of Mathematics, Statistics and Computer Science, University of Tabriz, Tabriz, Iran
*Center for Applied Optimization, Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA

ARTICLE INFO

Article history: Received 14 March 2022 Received in revised form 7 September 2022 Accepted 9 September 2022 Available online 20 September 2022

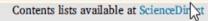
Keywords:
Wasserstein metric
Portfolio optimization
Semidefinite programming
Second-order stochastic dominance
Second-order conic programming
Ambiguity
Semi-infinite programming
Cutting surface method

ABSTRACT

In portfolio optimization, we may be dealing with misspecification of a known distribution, that stock returns follow it. The unknown true distribution is considered in terms of a Wasserstein-neighborhood of P to examine the tractable formulations of the portfolio selection problem. This study considers a distributionally robust portfolio optimization problem with an ambiguous stochastic dominance constraint by assuming the unknown distribution of asset returns. The objective is to maximize the worst-case expected return and subject to an ambiguous second-order stochastic dominance constraint. The expected return robustly stochastically dominates the benchmark in the second order over all possible distributions within an ambiguity set. It is also shown that the Wasserstein-moment ambiguity set-based distributionally robust portfolio optimization can be reduced to a semidefinite program and second-order conic programming. We use a cutting plane to solve our second-order stochastic dominance constraint portfolio optimization problem with ambiguity sets based on the Wasserstein metric. Then we decompose this class of distributionally robust portfolio optimization into semi-infinite programming and apply the cutting surface method to solve it. The captured optimization programs are applied to real-life data for more efficient comparison. The problems are examined in depth using the optimal solutions of the optimization programs based on the different setups.

© 2022 Elsevier Inc. All rights reserved.





Finance Research Letters

journal homepage: www.elsevier.com/locate/frl



Portfolio optimization using robust mean absolute deviation model: Wasserstein metric approach

Zohreh Hosseini-Nodeh a, Rashed Khanjani-Shiraz a,*, Panos M. Pardalos b

- a Faculty of Mathematics, Statistics and Computer Science, University of Tabriz, Tabriz, Iran
- ^b Center for Applied Optimization, Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA

ARTICLE INFO

Keywards:

Weighted mean absolute deviation Wasserstein metric Worst-case distribution Portfolio problem Robust optimization

ABSTRACT

Portfolio optimization can lead to misspecified stock returns that follow a known distribution. To investigate tractable formulations of the portfolio selection problem, we study these problems with the ambiguity set defined by the Wasserstein metric. Robust optimization with Wasserstein models protects against ambiguity in the distribution when analyzing decisions. This study considers portfolio optimization using a robust mean absolute deviation model consistent with the Wasserstein metric. The core of our idea is to consider the sets of distributions that lie within a certain distance from an empirical distribution. However, since information in financial markets is often unclear, we extend this structure to the weighted mean absolute deviation model when the underlying probability distribution is not precisely known. We then construct a decomposition algorithm based on the Benders decomposition approach to solve such problems. For more efficient comparison, the acquired optimization programs are applied to real data.

Future work

Ellipsoidal uncertainty set Third and N-order stochastic dominance

Application in portfolio optimization

Conclusion

- •Optimization is a powerful tool for decision-making in complex real-world environments.
- •We explored various types of optimization: linear, integer, nonlinear, stochastic, robust, and distributionally robust.
- Applications span across finance, energy, logistics, machine learning, and more.
- •Emphasis was placed on portfolio optimization under uncertainty using advanced optimization models.
- •My research contributes by integrating Distributionally Robust Optimization (DRO) and Wasserstein distance into risk-sensitive investment strategies.

Closing Statement:

Optimization not only enhances theoretical understanding but also delivers practical value across industries.

Questions & answers

Invite questions from the audience



Resources

List the resources you used for your research:

- Bertsimas, D., & Tsitsiklis, J. N. Introduction to Linear Optimization. Athena Scientific.
- Shapiro, A., Dentcheva, D., & Ruszczyński, A. Lectures on Stochastic Programming: Modeling and Theory.
- Ben-Tal, A., El Ghaoui, L., & Nemirovski, A. Robust Optimization. Princeton University Press.
- Esfahani, P. M., & Kuhn, D. (2018). Data-Driven Distributionally Robust Optimization Using the Wasserstein Metric. Operations Research.
- Hosseini-Nodeh Z, Khanjani-Shiraz R, Pardalos PM. Distributionally robust portfolio optimization with second-order stochastic dominance based on wasserstein metric. Information Sciences. 2022 Oct 1;613:828-52.
- Hosseini-Nodeh Z, Khanjani-Shiraz R, Pardalos PM. Portfolio optimization using robust mean absolute deviation model: Wasserstein metric approach. Finance Research Letters. 2023 Jun 1;54:103735.
- Shiraz RK, Nodeh ZH, Babapour-Azar A, Römer M, Pardalos PM.
 Distributionally robust joint chance-constrained programming:
 Wasserstein metric and second-order moment constraints. Information Sciences. 2024 Jan 1;654:119812.