Optimization with Constraint Learning LAB

Part I

Speaker: Donato Maragno
University of Amsterdam

February 8, 2023

Before starting

https://github.com/ordavidov/ocl_lab

Agenda

- Introduction to mathematical optimization
- Introduction to Constraint Learning
- Chemotherapy case study
- Embedding Predictive Models
- Trust region constraints
- Hands-on tutorial

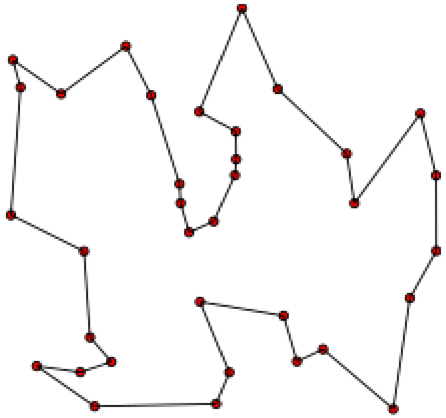
Mathematical Optimization

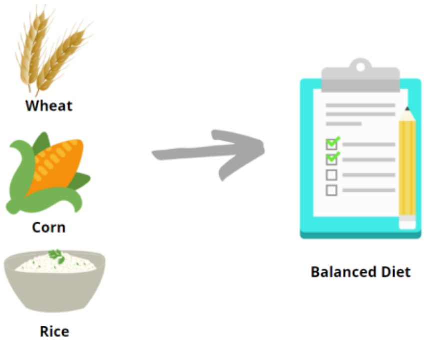
An optimization problem seeks to find the best (smallest or largest) value of a quantity given certain limits to the problem.

It can be expressed as a **minimization** (or maximization) of a quantity subject to a set of **constraints**.

Some examples:

- Travelling salesperson problem: minimize distance s.t. each city visited exactly once
- Diet problem: minimize cost s.t. nutrient requirements
- knapsack problem: maximize the value of objects in the knapsack s.t. capacity constraints





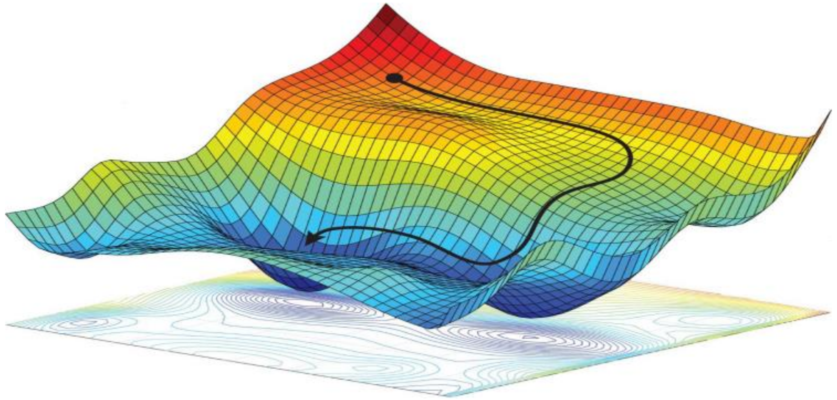
Mathematical Optimization

The general form of an optimization model is:

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \; f(x_1, \dots, x_n) \rightarrow \textit{Objective function}$$

$$\textit{subject to} \;\; g_i(x_1, \dots, x_n) \leq 0, \;\; i = 1, \dots, m \rightarrow \textit{Constraints}$$

where $x_1, \dots, x_n$ are the decision variables and the goal is to find a value for each of them such that the constraints are satisfied and the objective value is minimized.



Mathematical Optimization

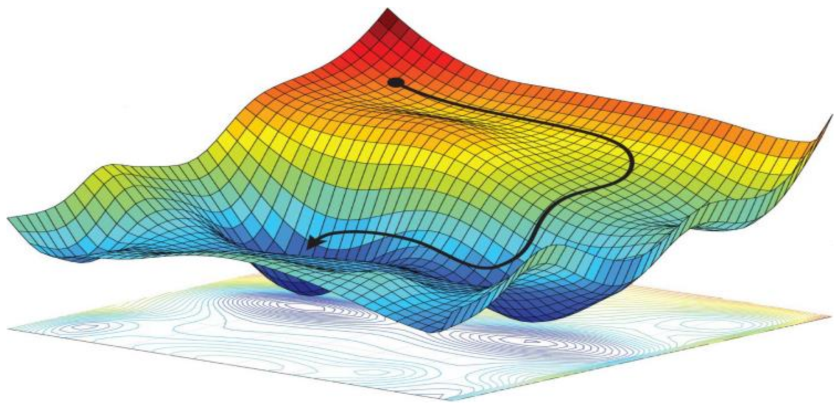
UNIVERSITY OF AMSTERDAM IBM

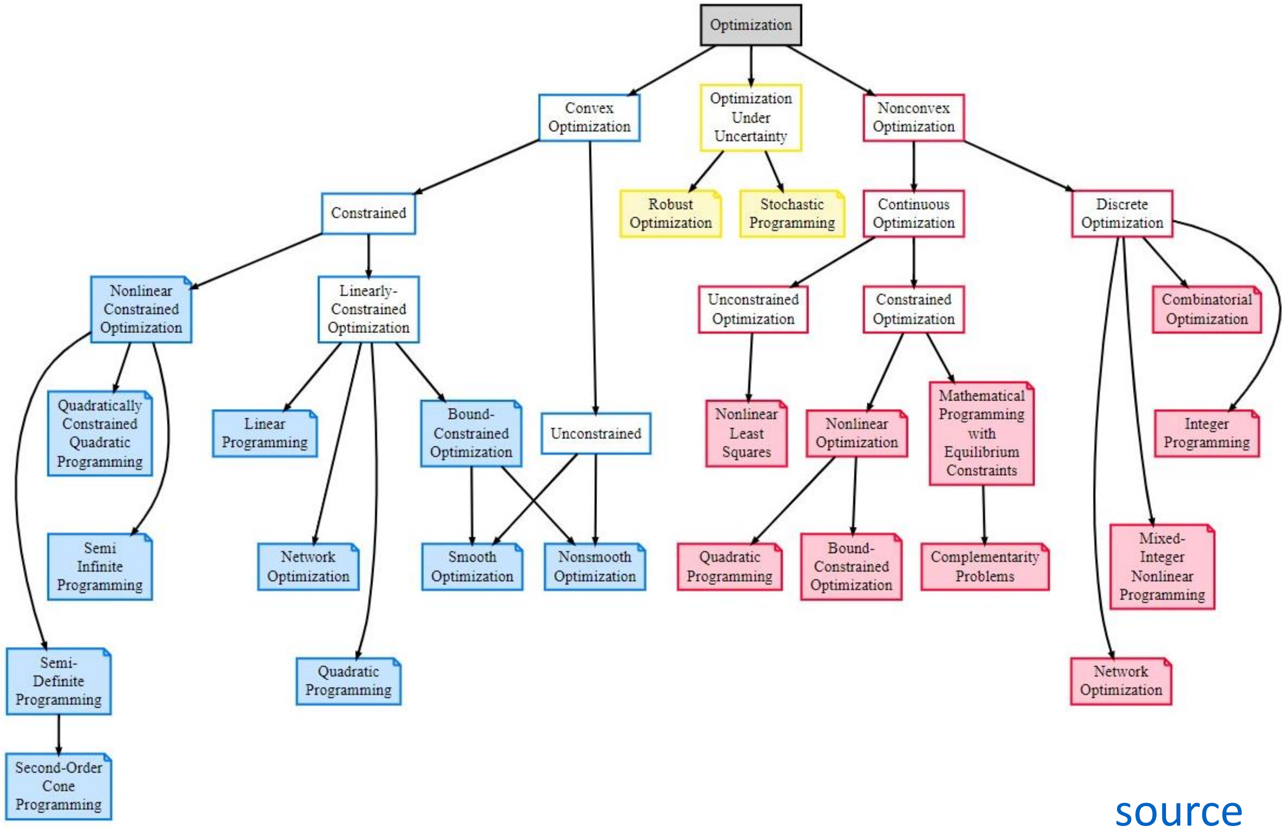
The general form of an optimization model is:

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} \ f(x_1, \dots, x_n) \rightarrow \textit{Objective function}$$

$$\textit{subject to} \quad g_i(x_1, \dots, x_n) \le 0, \ i = 1, \dots, m \rightarrow \textit{Constraints}$$

where $x_1, \dots, x_n$ are the decision variables and the goal is to find a value for each of them such that the constraints are satisfied and the objective value is minimized.





source

Constraint Learning

UNIVERSITY OF AMSTERDAM

Mixed-Integer Optimization (MIO)

Powerful tool that allows us to optimize a given objective subject to various constraints.

Many real-life optimization problems contain one or more constraints or objectives for which **there are no explicit formulae.**

Constraint Learning

UNIVERSITY OF AMSTERDAM IBM

Mixed-Integer Optimization (MIO)

Powerful tool that allows us to optimize a given objective subject to various constraints.

Many real-life optimization problems contain one or more constraints or objectives for which **there are no explicit formulae.**

Machine Learning (ML)

Data is available and machine learning models can be used to **learn the constraints.**

Case Study

Optimize the food baskets delivered by WFP to population groups in need.

- Minimize the procurement costs.
- Nutrient requirements constraints
- The food basket must be palatable

Case Study

Optimize the food baskets delivered by WFP to population groups in need.

- Minimize the procurement costs.
- Nutrient requirements constraints
- The food basket must be palatable

$$\underset{x \in \mathbb{R}^n_{\geq 0}}{\text{minimize}} \quad f(\boldsymbol{x}) \longrightarrow \textit{Procurement costs}$$

Beans	Fish	...	Oil

Case Study

Optimize the food baskets delivered by WFP to population groups in need.

- Minimize the procurement costs.
- Nutrient requirements constraints
- The food basket must be palatable

Energy	Fat	...	Vitamin C

$$\underset{\boldsymbol{x}\in\mathbb{R}^n_{\geq 0}}{minimize} \quad f(\boldsymbol{x})$$ *Procurement costs*

$$subject\ to \quad \boldsymbol{NutVal}_i^T \boldsymbol{x} \geq NutReq_i \quad \forall\, i \in \mathcal{N}$$ *Nutrient requirements constraints*

Beans	Fish	...	Oil

Case Study

WFP

UNIVERSITY OF AMSTERDAM

IBM

Optimize the food baskets delivered by WFP to population groups in need.

- Minimize the procurement costs.
- Nutrient requirements constraints
- The food basket must be palatable

Energy	Fat	...	Vitamin C

$$\begin{aligned} &\underset{x \in \mathbb{R}^n_{\geq 0}}{minimize} && f(\boldsymbol{x}) \\ &subject\ to && \boldsymbol{NutVal}_i^T \boldsymbol{x} \geq NutReq_i \quad \forall\, i \in \mathcal{N} \\ & && Palatability(\boldsymbol{x}) \geq t \end{aligned}$$

Procurement costs

Nutrient requirements constraints

No explicit formula but data is available

Beans	Fish	...	Oil

Applications of Constraint Learning

Table 1: Methods used for constraint learning

	Neural Networks	Decision Trees	Random Forest	Other Ensemble	Support Vector Machines	Clustering	(M)ILP	Other
Bergman et al. (2019)	x							x
Biggs et al. (2018)		x	x					
Chen et al. (2020)	x							
Chi et al. (2007)					x			
Cozad et al. (2014)							x	x
Cremer et al. (2018)		x		x				
De Angelis et al. (2003)	x							
Fahmi and Cremaschi (2012)	x							x
Garg et al. (2018)					x			
Grimstad and Andersson (2019)	x							
Gutierrez-Martinez et al. (2010)	x							
Halilbašić et al. (2018)		x						
Jalali et al. (2019)					x			
Kudła and Pawlak (2018)		x						
Lombardi et al. (2017)	x	x						
Maragno et al. (2022)	x	x	x	x	x	x		x
Mišić (2020)		x	x					
Paulus et al. (2021)	x							
Pawlak and Krawiec (2017a)							x	
Pawlak and Krawiec (2017b)								x
Pawlak and Krawiec (2018)								x
Pawlak (2019)								x
Pawlak and Litwiniuk (2021)						x		x
Pawlak and O'Neill (2021)							x	x
Prat and Chatzivasileiadis (2020)		x						
Say et al. (2017)	x							
Schede et al. (2019)		x					x	
Schweidtmann and Mitsos (2019)	x							
Spyros (2020)	x	x						
Sroka and Pawlak (2018)						x		x
Thams et al. (2017)		x						
Venzke et al. (2020b)	x							
Verwer et al. (2017)								
Xavier et al. (2021)								
Yang and Bequette (2021)	x							

Fajemisin A, Maragno D, den Hertog D (2021) Optimization with constraint learning: A framework and survey. URL https://arxiv.org/abs/2110.02121.

Applications of Constraint Learning

UNIVERSITY OF AMSTERDAM

Table 1: Methods used for constraint learning

	Neural Networks	Decision Trees	Random Forest	Other Ensemble	Support Vector Machines	Clustering	(M)ILP	Other
Bergman et al. (2019)	x							x
Biggs et al. (2018)		x	x					
Chen et al. (2020)	x							
Chi et al. (2007)					x			
Cozad et al. (2014)							x	x
Cremer et al. (2018)		x		x				
De Angelis et al. (2003)	x							
Fahmi and Cremaschi (2012)	x							x
Garg et al. (2018)					x			
Grimstad and Andersson (2019)	x							
Gutierrez-Martinez et al. (2010)	x							
Halilbašić et al. (2018)		x						
Jalali et al. (2019)					x			
Kudła and Pawlak (2018)		x						
Lombardi et al. (2017)	x	x						
Maragno et al. (2022)	x	x	x	x	x	x		x
Mišić (2020)		x	x					
Paulus et al. (2021)	x							
Pawlak and Krawiec (2017a)							x	
Pawlak and Krawiec (2017b)								x
Pawlak and Krawiec (2018)								x
Pawlak (2019)								x
Pawlak and Litwiniuk (2021)						x		x
Pawlak and O'Neill (2021)							x	x
Prat and Chatzivasileiadis (2020)		x						
Say et al. (2017)	x							
Schede et al. (2019)		x					x	
Schweidtmann and Mitsos (2019)	x							
Spyros (2020)	x	x						
Sroka and Pawlak (2018)						x		x
Thams et al. (2017)		x						
Venzke et al. (2020b)	x							
Verwer et al. (2017)								
Xavier et al. (2021)								
Yang and Bequette (2021)	x							

Say, B., Wu, G., Zhou, Y.Q., Sanner, S., 2017. Nonlinear hybrid planning with deep net learned transition models and mixed-integer linear programming, in: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence IJCAI-17, pp. 750–756.

Optimal planning using mixed integer linear optimization and a fitted deep network transition model

Fajemisin A, Maragno D, den Hertog D (2021) Optimization with constraint learning: A framework and survey. URL https://arxiv.org/abs/2110.02121.

Applications of Constraint Learning

Table 1: Methods used for constraint learning

	Neural Networks	Decision Trees	Random Forest	Other Ensemble	Support Vector Machines	Clustering	(M)ILP	Other
Bergman et al. (2019)	x							x
Biggs et al. (2018)		x	x					
Chen et al. (2020)	x							
Chi et al. (2007)					x			
Cozad et al. (2014)							x	x
Cremer et al. (2018)		x		x				
De Angelis et al. (2003)	x							
Fahmi and Cremaschi (2012)	x							x
Garg et al. (2018)					x			
Grimstad and Andersson (2019)	x							
Gutierrez-Martinez et al. (2010)	x							
Halilbašić et al. (2018)		x						
Jalali et al. (2019)					x			
Kudła and Pawlak (2018)		x						
Lombardi et al. (2017)	x	x						
Maragno et al. (2022)	x	x	x	x	x	x		
Mišić (2020)		x	x					
Paulus et al. (2021)	x							
Pawlak and Krawiec (2017a)							x	
Pawlak and Krawiec (2017b)								
Pawlak and Krawiec (2018)								
Pawlak (2019)								x
Pawlak and Litwiniuk (2021)						x		x
Pawlak and O'Neill (2021)							x	x
Prat and Chatzivasileiadis (2020)		x						
Say et al. (2017)	x							
Schede et al. (2019)		x					x	
Schweidtmann and Mitsos (2019)	x							
Spyros (2020)	x	x						
Sroka and Pawlak (2018)						x		x
Thams et al. (2017)		x						
Venzke et al. (2020b)	x							
Verwer et al. (2017)								
Xavier et al. (2021)								
Yang and Bequette (2021)	x							

Say, B., Wu, G., Zhou, Y.Q., Sanner, S., 2017. Nonlinear hybrid planning with deep net learned transition models and mixed-integer linear programming, in: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence IJCAI-17, pp. 750–756.

Optimal planning using mixed integer linear optimization and a fitted deep network transition model

Global Optimization via Optimal Decision Trees

Dimitris Bertsimas

Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA, dbertsim@mit.edu

Berk Öztürk

Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA, bozturk@mit.edu

Approximate non-convex function with Decision trees

Fajemisin A, Maragno D, den Hertog D (2021) Optimization with constraint learning: A framework and survey. URL https://arxiv.org/abs/2110.02121.

Constraint Learning

$\boldsymbol{x}$ Decision variables

$\boldsymbol{w}$ Contextual variables

$\boldsymbol{y}$ Predicted outcomes

Constraint Learning

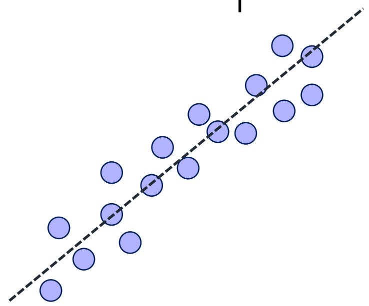
$\boldsymbol{x}$ Decision variables

$\boldsymbol{w}$ Contextual variables

$\boldsymbol{y}$ Predicted outcomes

$$\boldsymbol{y} = \hat{\boldsymbol{h}}_{\mathcal{D}}(\boldsymbol{x}, \boldsymbol{w})$$

$$D = \{(\overline{\boldsymbol{x}}_i, \overline{\boldsymbol{w}}_i, \overline{\boldsymbol{y}}_i)\}_{i=1}^{N} \longrightarrow$$



Constraint Learning

$$\min_{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{y}\in\mathbb{R}^k} f(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y})$$

$$\text{s.t. } \boldsymbol{g}(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y}) \leq \boldsymbol{0}$$

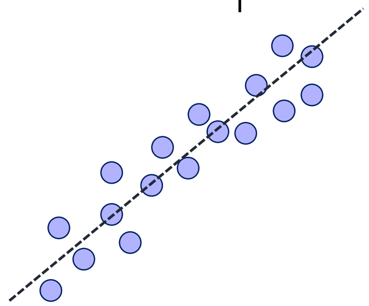
$$\boxed{\boldsymbol{y} = \hat{\boldsymbol{h}}_{\mathcal{D}}(\boldsymbol{x},\boldsymbol{w})}$$

$\boldsymbol{x}$ Decision variables

$\boldsymbol{w}$ Contextual variables

$\boldsymbol{y}$ Predicted outcomes

$$D = \{(\overline{\boldsymbol{x}}_i, \overline{\boldsymbol{w}}_i, \overline{\boldsymbol{y}}_i)\}_{i=1}^{N} \longrightarrow$$



Constraint Learning

$$\min_{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{y}\in\mathbb{R}^k} \quad f(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y})$$

$$\text{s.t. } \boldsymbol{g}(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y}) \le \boldsymbol{0}$$

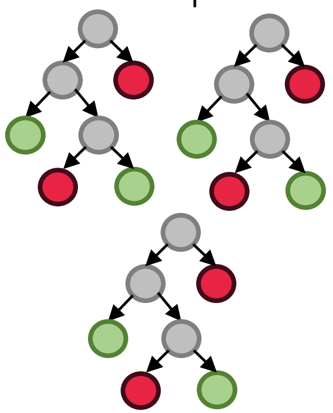
$$\boxed{\boldsymbol{y} = \hat{\boldsymbol{h}}_{\mathcal{D}}(\boldsymbol{x},\boldsymbol{w})}$$

$\boldsymbol{x}$ Decision variables

$\boldsymbol{w}$ Contextual variables

$\boldsymbol{y}$ Predicted outcomes

$$D = \{(\overline{\boldsymbol{x}}_i, \overline{\boldsymbol{w}}_i, \overline{\boldsymbol{y}}_i)\}_{i=1}^{N} \longrightarrow$$



Constraint Learning

$$\min_{\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{y}\in\mathbb{R}^k} \quad f(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y})$$

$\boldsymbol{x}$ Decision variables

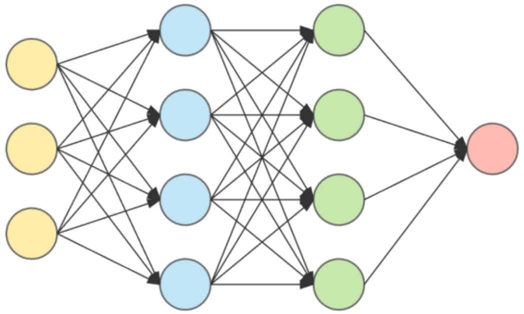
$\boldsymbol{w}$ Contextual variables

$\boldsymbol{y}$ Predicted outcomes

$$\text{s.t. } \boldsymbol{g}(\boldsymbol{x},\boldsymbol{w},\boldsymbol{y}) \leq \boldsymbol{0}$$

$$\boxed{\boldsymbol{y} = \hat{\boldsymbol{h}}_{\mathcal{D}}(\boldsymbol{x},\boldsymbol{w})}$$

$$D = \{(\overline{\boldsymbol{x}}_i, \overline{\boldsymbol{w}}_i, \overline{\boldsymbol{y}}_i)\}_{i=1}^{N} \longrightarrow$$



Chemotherapy Case Study

In this case study, we extend the work of Bertsimas et al. (2016)* in the design of chemotherapy regimens for advanced gastric cancer. Given a new study cohort and study characteristics, we would like to optimize a chemotherapy regimen to **maximize the cohort's survival subject to constraint on different types of toxicity**.

$$\begin{aligned} \boldsymbol{x}_b^d &= \mathbb{I}(\text{drug } d \text{ is administered}), \\ \boldsymbol{x}_a^d &= \text{average daily dose of drug } d, \\ \boldsymbol{x}_i^d &= \text{maximum instantaneous dose of drug } d. \end{aligned}$$

$$\boxed{\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{y}} \quad & y_{OS} \\ \text{s.t.}\, & y_i \le \tau_i, && i \in \mathcal{Y}_C, \\ & y_i = \hat{h}_i(\boldsymbol{x}, \boldsymbol{w}), && i \in \mathcal{Y}_C, \\ & y_{OS} = \hat{h}_{OS}(\boldsymbol{x}, \boldsymbol{w}), \end{aligned}}$$

$$\begin{aligned} & \sum_d \boldsymbol{x}_b^d \le 3, \\ & \boldsymbol{x}_b \in \{0,1\}^d, \\ & \boldsymbol{x} \in \mathcal{X}(\boldsymbol{w}). \end{aligned}$$

(*) Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. Management Science 62(5):1511–1531, ISSN 15265501

Chemotherapy Case Study

In this case study, we extend the work of [Bertsimas et al. (2016)]* in the design of chemotherapy regimens for advanced gastric cancer. Given a new study cohort and study characteristics, we would like to optimize a chemotherapy regimen to **maximize the cohort's survival subject to constraint on different types of toxicity**.

$$
\begin{aligned}
x_b^d &= \mathbb{I}(\text{drug } d \text{ is administered}),\\
x_a^d &= \text{average daily dose of drug } d,\\
x_i^d &= \text{maximum instantaneous dose of drug } d.
\end{aligned}
$$

$$
\begin{aligned}
\boxed{\begin{aligned}
\min_{x,y}\;& y_{OS} \\
\text{s.t.}\; & y_i \le \tau_i, && i \in \mathcal{Y}_C,\\
& y_i = \hat{h}_i(x, w), && i \in \mathcal{Y}_C,\\
& y_{OS} = \hat{h}_{OS}(x, w),
\end{aligned}}\\
& \sum_d x_b^d \le 3,\\
& x_b \in \{0,1\}^d,\\
& x \in \mathcal{X}(w).
\end{aligned}
$$

Cohort contextual variables

- Gender
- Age
- primary site breakdown
- ecog score
- …

(*) Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. Management Science 62(5):1511–1531, ISSN 15265501

Chemotherapy Case Study

UNIVERSITY OF AMSTERDAM IBM

In this case study, we extend the work of Bertsimas et al. (2016)* in the design of chemotherapy regimens for advanced gastric cancer. Given a new study cohort and study characteristics, we would like to optimize a chemotherapy regimen to **maximize the cohort's survival subject to constraint on different types of toxicity**.

$$\boldsymbol{x}_b^d = \mathbb{I}(\text{drug } d \text{ is administered}),$$
$$\boldsymbol{x}_a^d = \text{average daily dose of drug } d,$$
$$\boldsymbol{x}_i^d = \text{maximum instantaneous dose of drug } d.$$

Overall Survival

$$\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{y}} \quad & y_{OS} \\
\text{s.t.}\quad & y_i \leq \tau_i, && i \in \mathcal{Y}_C, \\
& y_i = \hat{h}_i(\boldsymbol{x}, \boldsymbol{w}), && i \in \mathcal{Y}_C, \\
& y_{OS} = \hat{h}_{OS}(\boldsymbol{x}, \boldsymbol{w}), \\
& \sum_d \boldsymbol{x}_b^d \leq 3, \\
& \boldsymbol{x}_b \in \{0,1\}^d, \\
& \boldsymbol{x} \in \mathcal{X}(\boldsymbol{w}).
\end{aligned}$$

Cohort contextual variables

- Gender
- Age
- primary site breakdown
- ecog score
- …

(*) Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. Management Science 62(5):1511–1531, ISSN 15265501

Chemotherapy Case Study

In this case study, we extend the work of [Bertsimas et al. (2016)]* in the design of chemotherapy regimens for advanced gastric cancer. Given a new study cohort and study characteristics, we would like to optimize a chemotherapy regimen to **maximize the cohort's survival subject to constraint on different types of toxicity**.

$$\begin{aligned} \boldsymbol{x}_b^d &= \mathbb{I}(\text{drug } d \text{ is administered}),\\ \boldsymbol{x}_a^d &= \text{average daily dose of drug } d,\\ \boldsymbol{x}_i^d &= \text{maximum instantaneous dose of drug } d. \end{aligned}$$

Overall Survival

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{y}} \quad & y_{OS} \\ \text{s.t.}\; & y_i \le \tau_i, && i \in \mathcal{Y}_C, \\ & y_i = \hat{h}_i(\boldsymbol{x}, \boldsymbol{w}), && i \in \mathcal{Y}_C, \\ & y_{OS} = \hat{h}_{OS}(\boldsymbol{x}, \boldsymbol{w}), \\ & \sum_d \boldsymbol{x}_b^d \le 3, \\ & \boldsymbol{x}_b \in \{0,1\}^d, \\ & \boldsymbol{x} \in \mathcal{X}(\boldsymbol{w}). \end{aligned}$$

Toxicities

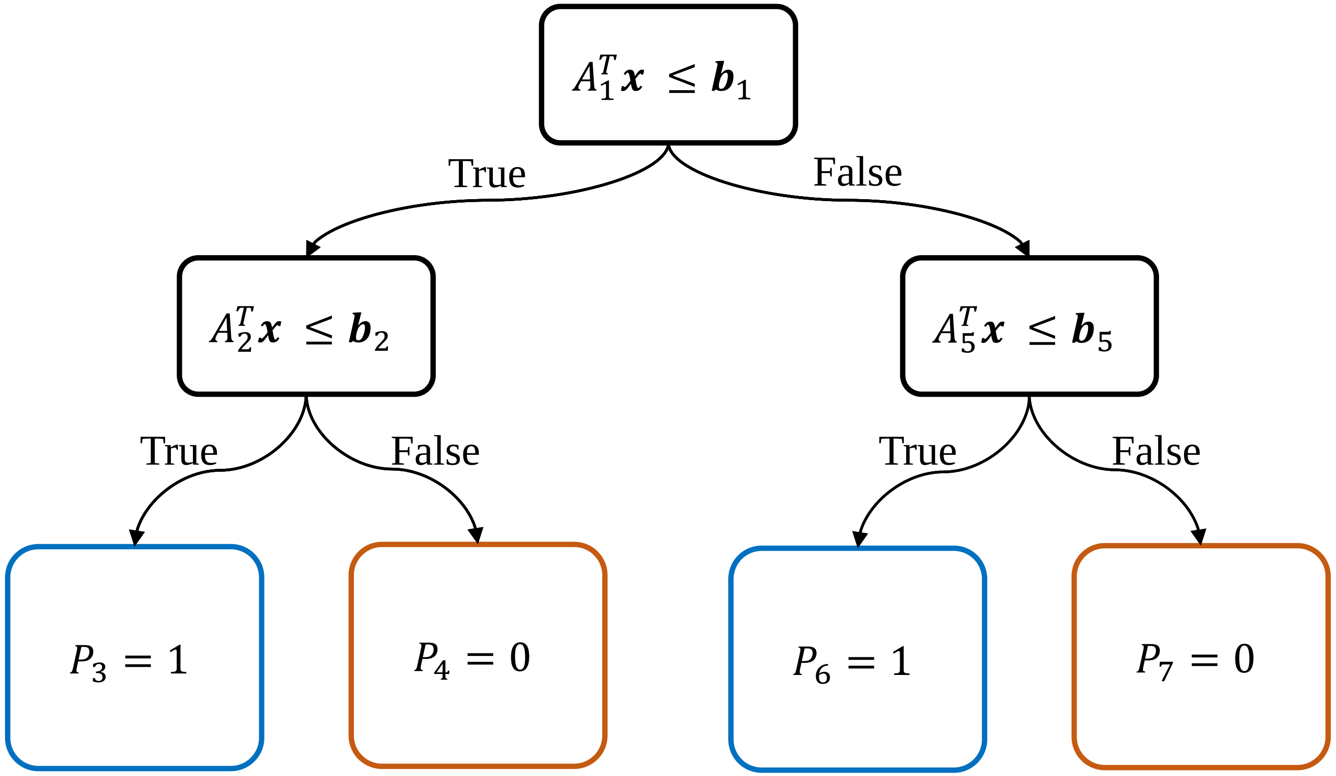
- Grade 3/4 constitutional
- Infection
- Neurological
- Grade 4 blood
- …

Cohort contextual variables

- Gender
- Age
- primary site breakdown
- ecog score
- …

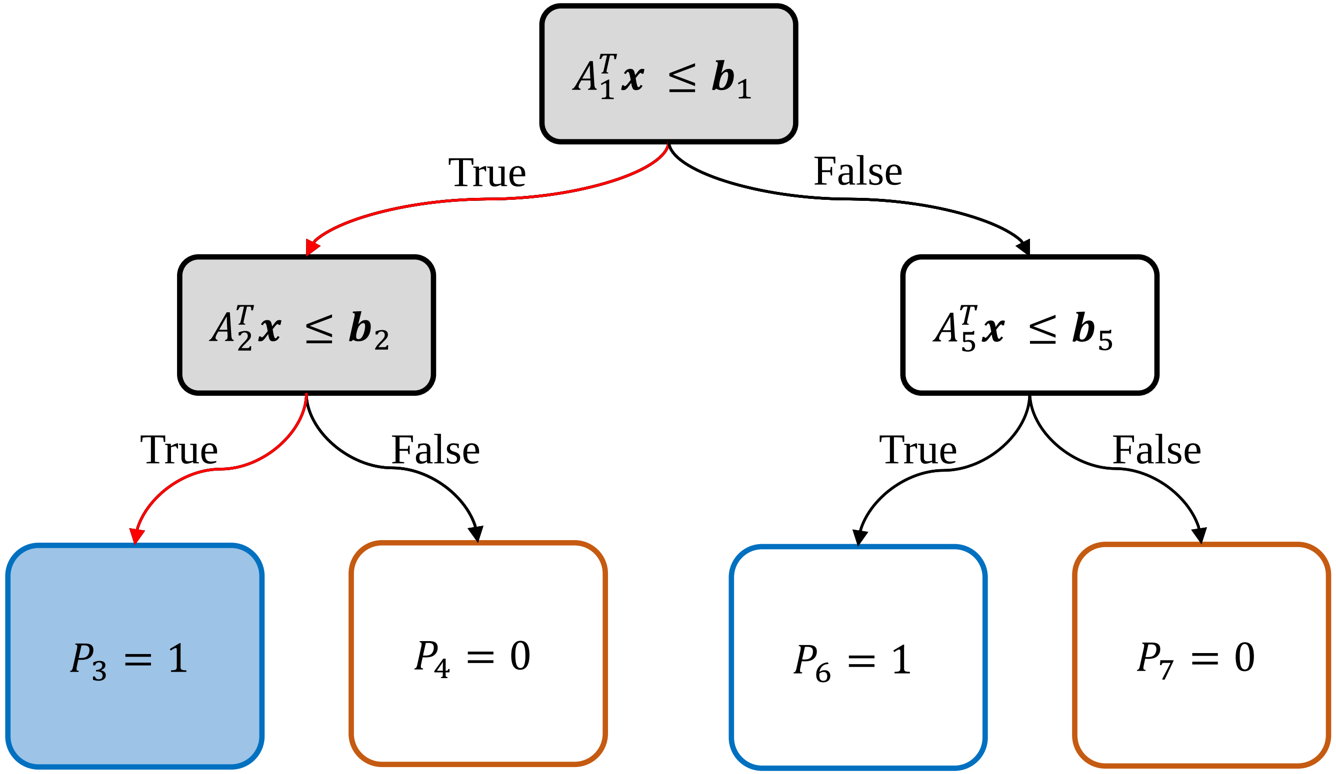
(*) Bertsimas D, O'Hair A, Relyea S, Silberholz J (2016) An analytics approach to designing combination chemotherapy regimens for cancer. Management Science 62(5):1511–1531, ISSN 15265501

Embedding Decision Trees



Embedding Decision Trees

UNIVERSITY OF AMSTERDAM IBM



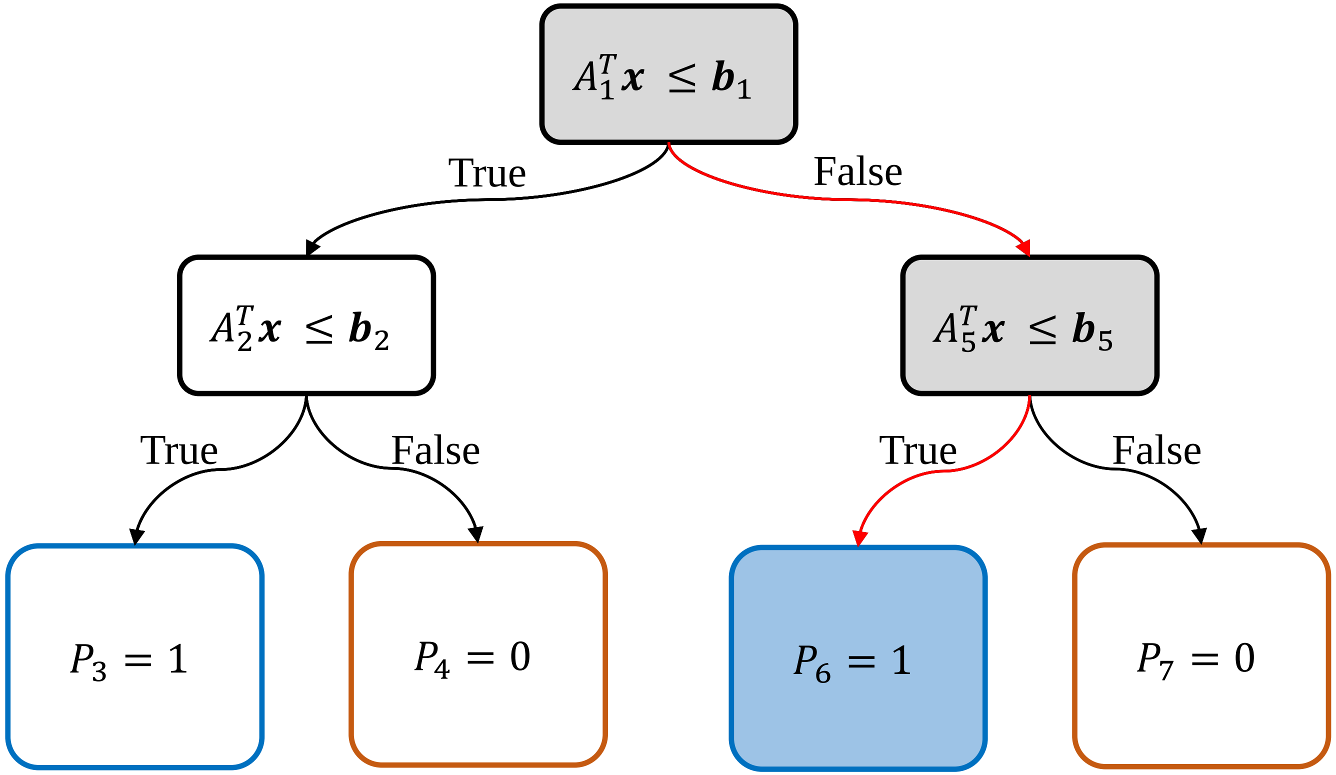
Conceptual model

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0} \end{aligned}$$

P_3

$$A_1^T \boldsymbol{x} \le \boldsymbol{b}_1$$
$$A_2^T \boldsymbol{x} \le \boldsymbol{b}_2$$

Embedding Decision Trees



Conceptual model

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \end{aligned}$$

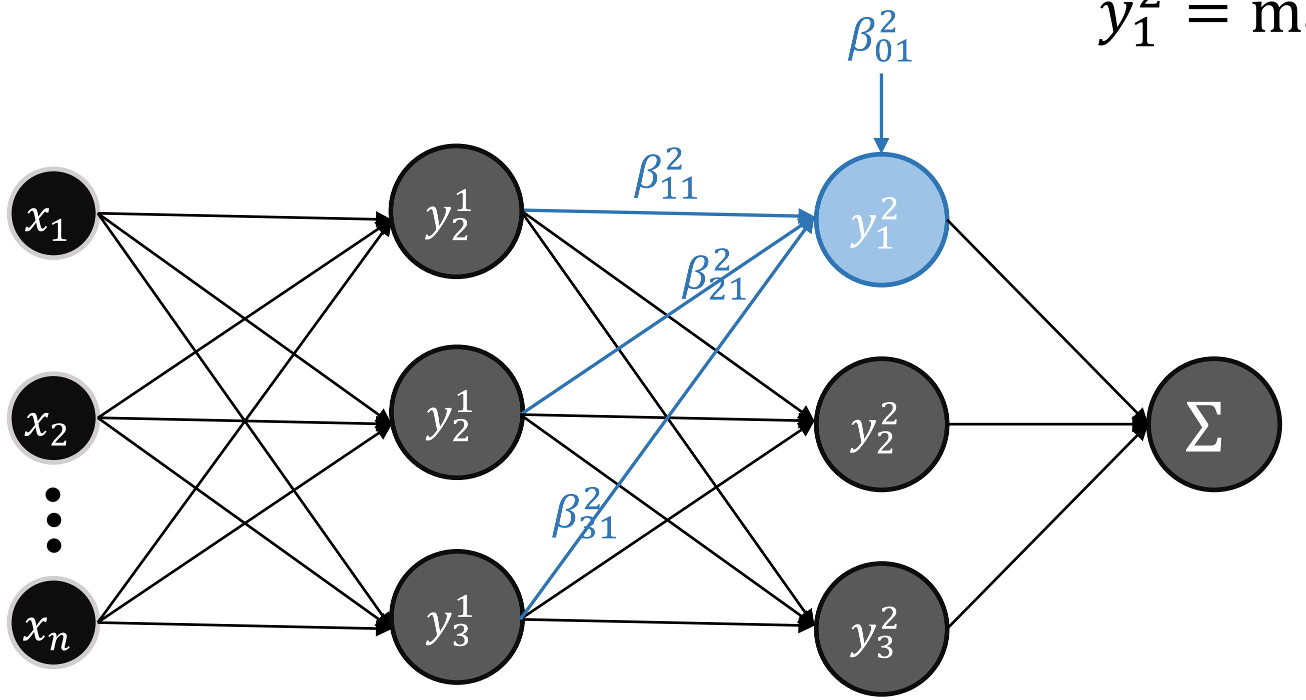
P_6

$$A_1^T\boldsymbol{x} \geq \boldsymbol{b}_1$$
$$A_5^T\boldsymbol{x} \leq \boldsymbol{b}_5$$

Embedding Neural Networks

ReLU activation function

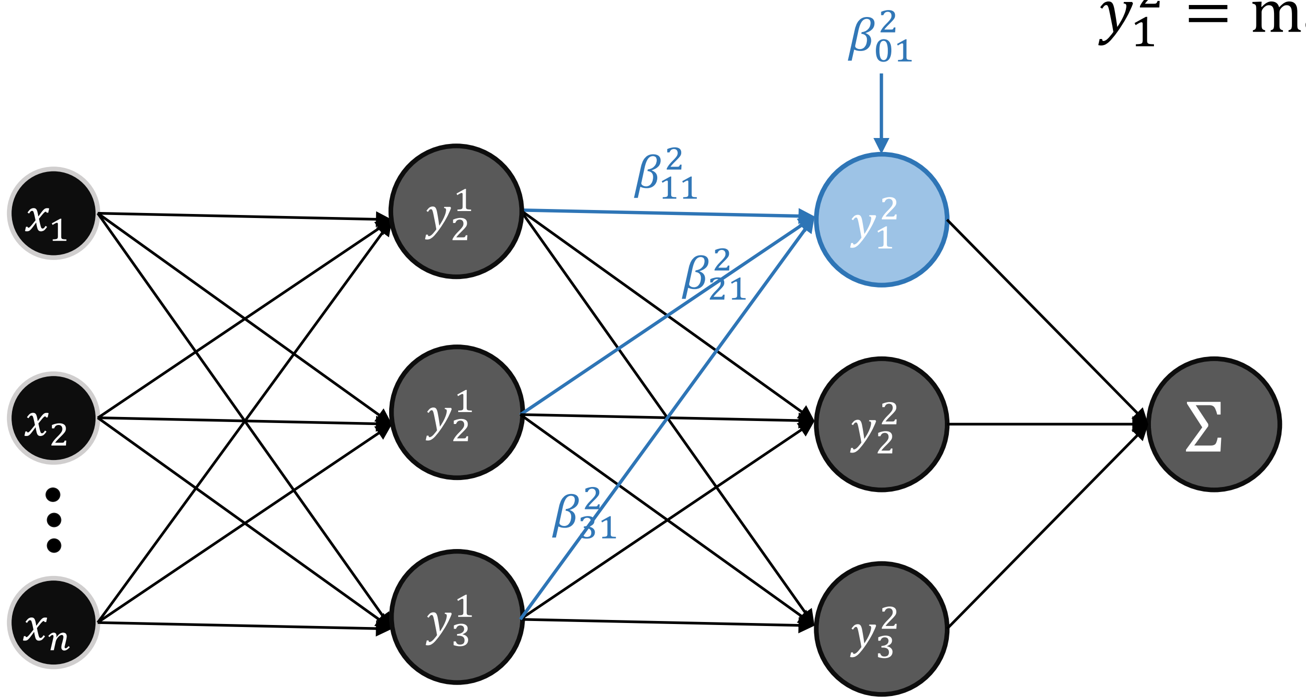
$$y_1^2 = \max\left\{0, \beta_{01}^2 + \boldsymbol{\beta}_1^{2^T}\boldsymbol{y}^1\right\}$$



Embedding Neural Networks

ReLU activation function

$$y_1^2 = \max\left\{0, \beta_{01}^2 + \boldsymbol{\beta}_1^{2^T} \boldsymbol{y}^1\right\}$$

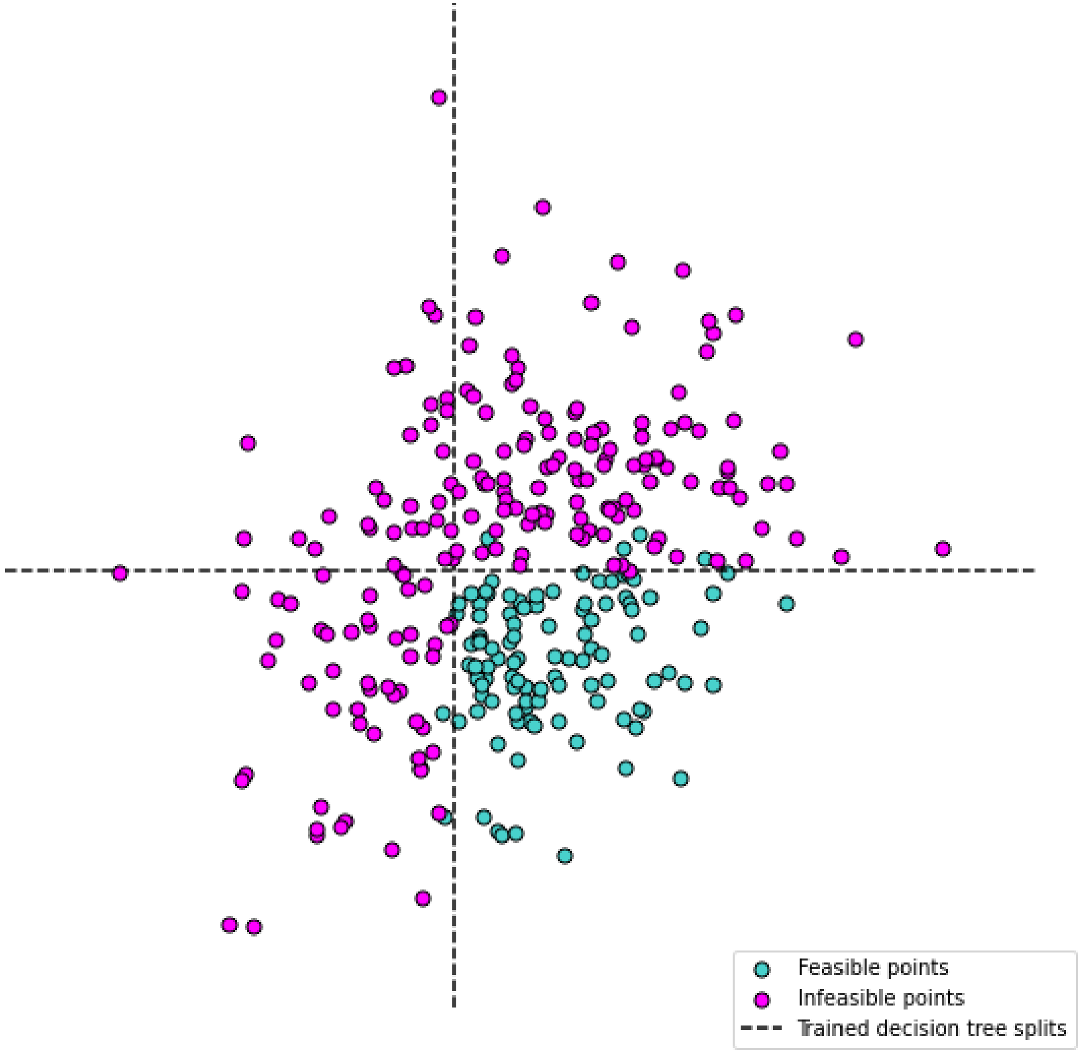


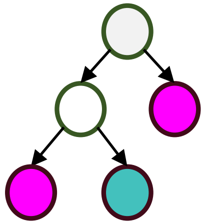
$y = \max\{0, x\}$ can also be written as:

$$\begin{cases} y \geq x, \\ y \leq x - M_L(1-z), \\ y \leq M_U z, \\ y \geq 0, \\ z \in \{0,1\}, \end{cases}$$

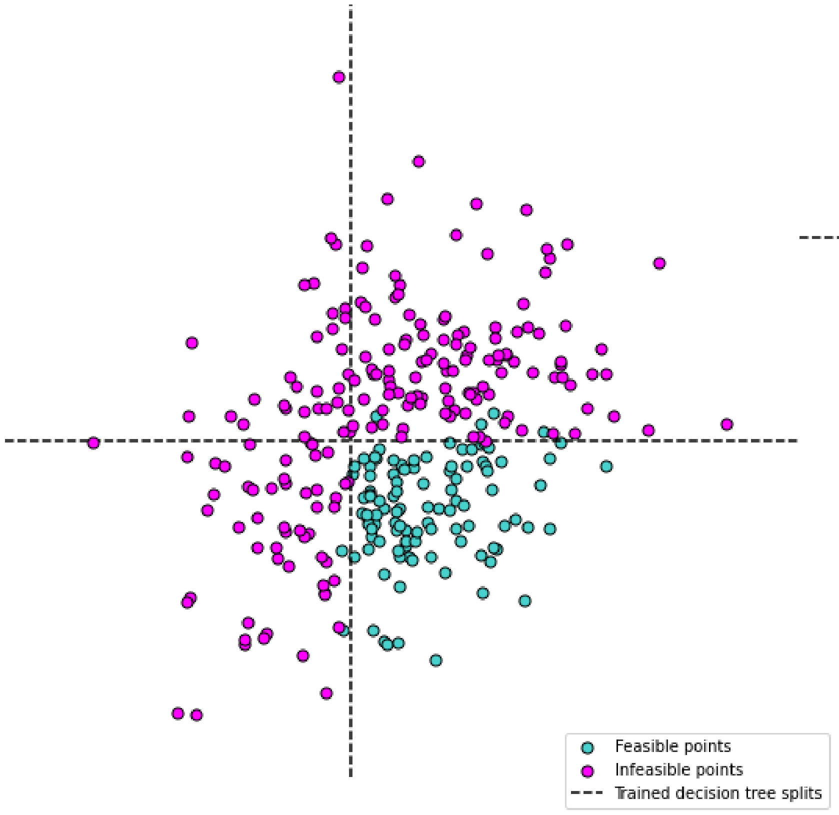
where $M_L < 0$ is a lower bound on all possible values of x, and $M_U > 0$ is an upper bound.

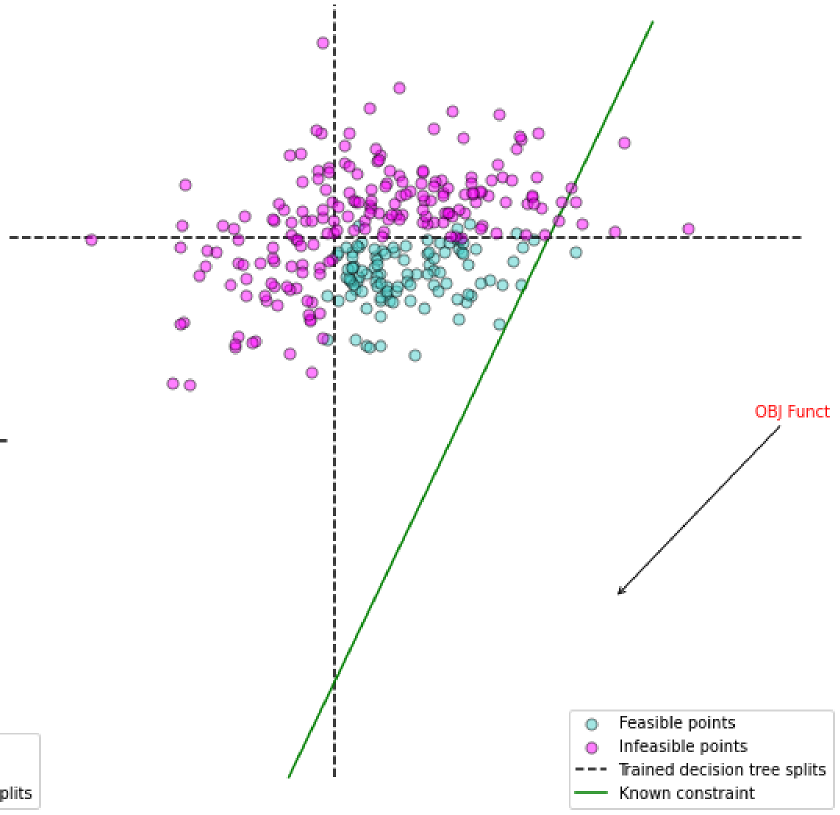
Trust region constraints

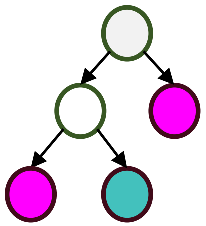




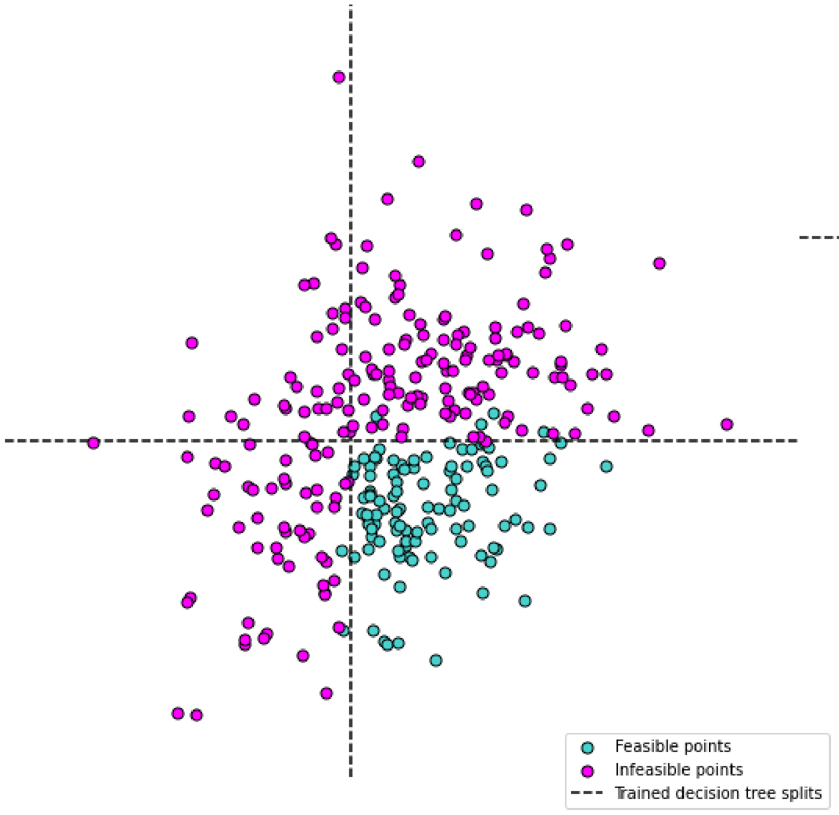
Trust region constraints

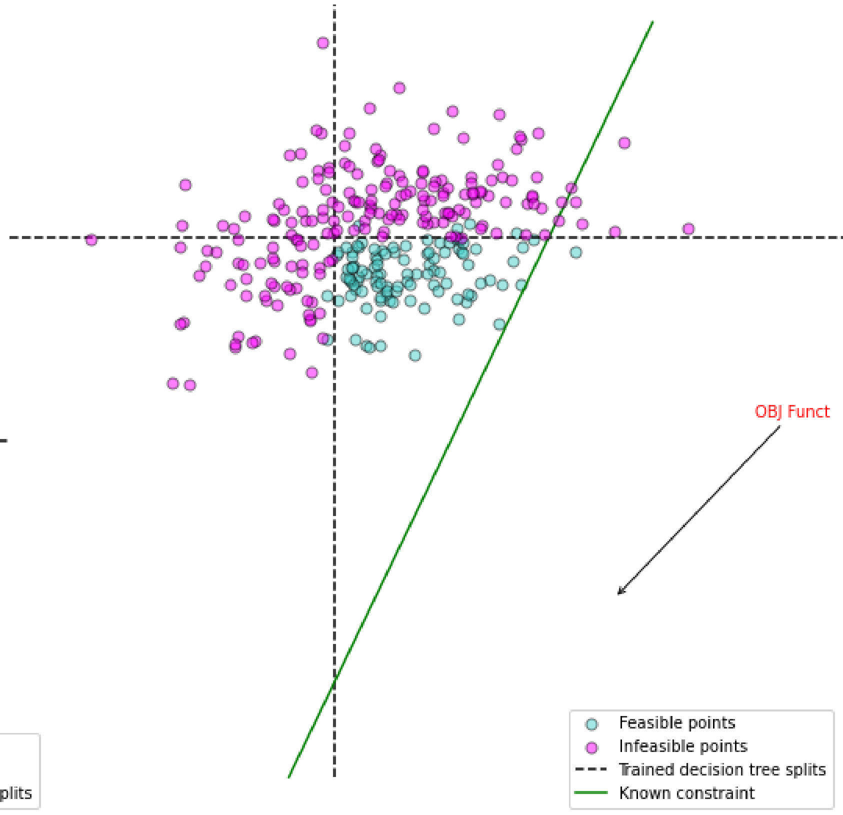


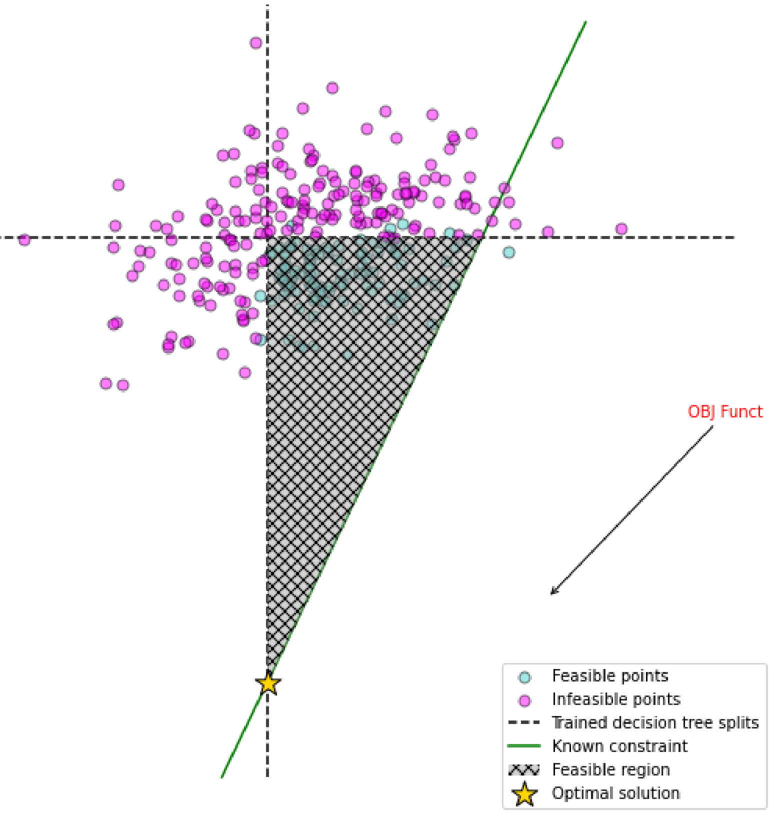


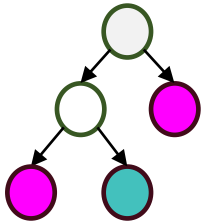


Trust region constraints

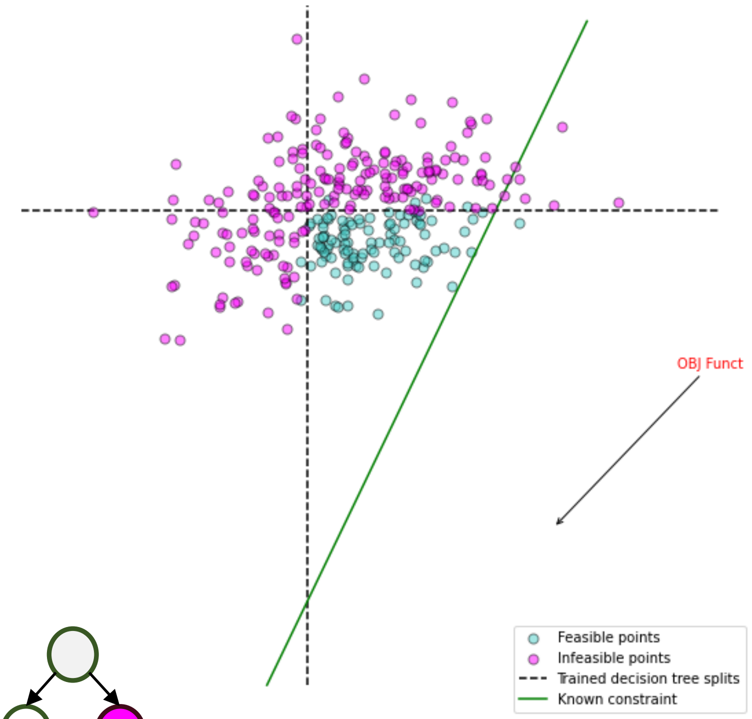


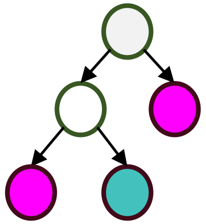






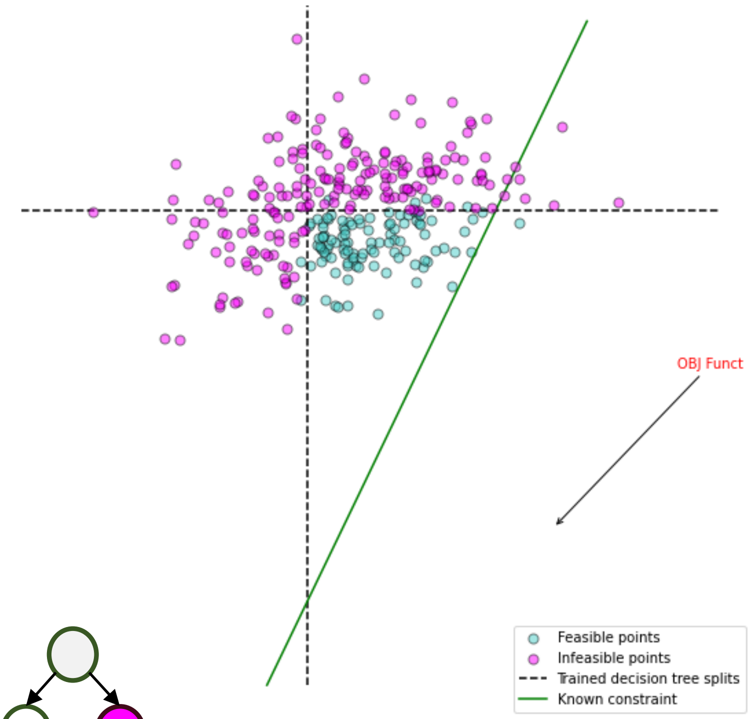
Trust region constraints

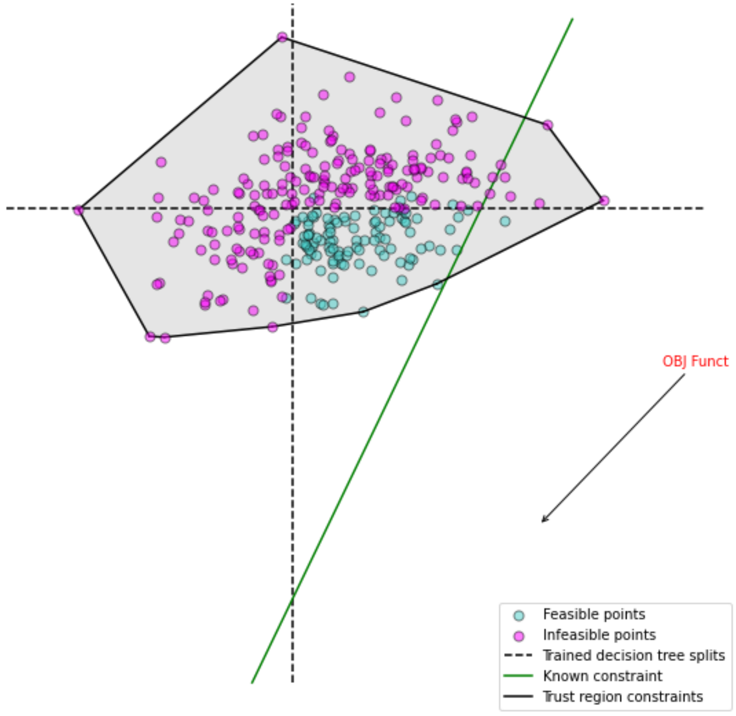


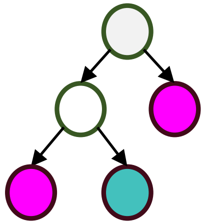


Trust region constraints

$$CH(\boldsymbol{x}) = \left\{ \boldsymbol{x} \,|\, \boldsymbol{x} = \sum_{i}^{N} \lambda_i \bar{\boldsymbol{x}}_i \,, \sum_{i}^{N} \lambda_i = 1 \,, \lambda_i \geq 0, i = 1, \ldots, N \right\}$$

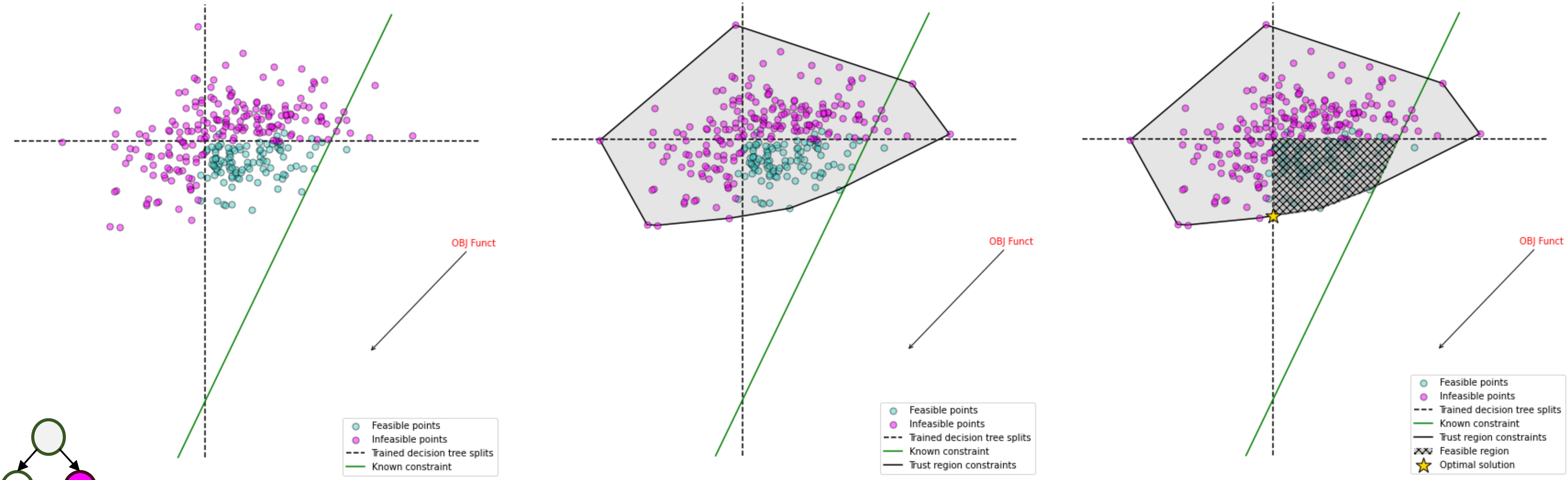


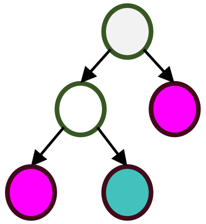




Trust region constraints

$$CH(\boldsymbol{x}) = \left\{ \boldsymbol{x} | \boldsymbol{x} = \sum_{i}^{N} \lambda_i \overline{\boldsymbol{x}}_i \,, \sum_{i}^{N} \lambda_i = 1 \,, \lambda_i \geq 0, i = 1, \dots, N \right\}$$





OptiCL

A Python Package for Optimization with Constraint Learning

https://github.com/hwiberg/OptiCL

OptiCL

A Python Package for

Optimization with **C**onstraint **L**earning

https://github.com/hwiberg/OptiCL

Hands-on tutorial on the

diet problem

Thank you!

Q&A

CONTACT ME

d.maragno@uva.nl

github.com/donato-maragno

@MaragnoDonato

linkedin.com/in/donato-maragno