Prof. Dr.-Ing. habil. S. Mostaghim
Faculty of Computer Science
Otto von Guericke University Magdeburg

# Tutorials for EMO Lecture Summer Semester 2024

## Assignment 1 Meta-heuristic Optimization and Pareto Dominance

Please answer the following questions regarding meta-heuristic optimization and Pareto-dominance.

- What is the difference between a heuristic and a meta-heuristic?
- Why do we use meta-heuristic methods like Evolutionary Algorithms instead of classical, analytical methods?
- What is the difference between a single- and a multi-objective optimization problem?
- Explain the concept of Pareto-dominance and why it is used in multi-objective optimization.
- Consider the car example from the lecture. You want to buy a new car and have three criteria: Price (which should be minimized), gas consumption (which should me minimized) and maximum speed (which should be maximized). The cars and their properties are listed as follows.

| | Price (EUR) | Fuel consumption | Max speed (km/h) |
|---|---|---|---|
| VW | 16,200 | 7.2 | 180 |
| Opel | 14,900 | 7.0 | 220 |
| Ford | 14,000 | 7.5 | 200 |
| Toyota | 15,200 | 8.2 | 250 |

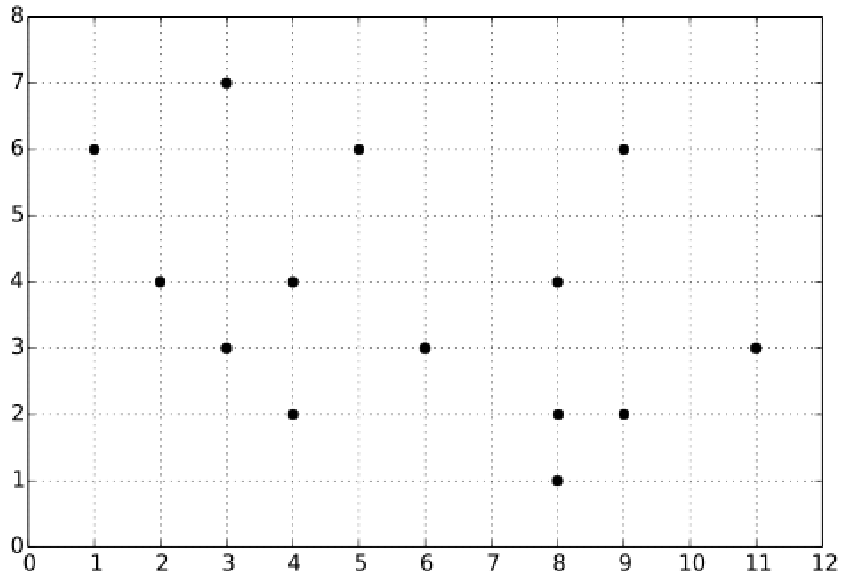
  Using Pareto-dominance, which of these cars are non-dominated? Which of these cars would you buy as a decision maker and why?

**Assignment 2 Pareto-Dominance Concepts**

Please answer the following questions regarding Pareto-dominance.

- What are the ideal points and nadir points in the context of multi-objective optimization?
- The following picture shows a set of solutions in a 2-dimensional objective space. ($f_1(\vec{x})$ on the $x$-axis and $f_2(\vec{x})$ on the $y$-axis). Indicate which of these solutions are non-dominanted with regard to Pareto-optimality if

  a) $f_1$ is minimized and $f_2$ is minimized
  b) $f_1$ is minimized and $f_2$ is maximized
  c) $f_1$ is maximized and $f_2$ is minimized
  d) $f_1$ is maximized and $f_2$ is maximized




Are the resultant non-dominated sets convex or concave? What are the ideal and nadir points of these sets (assuming there are no other solutions in the search space)? Indicate which solutions are non-dominated with regard to weak Pareto-optimality.

**Assignment 3**

We are going this holiday on a trip to Munich, and we are looking for a good hotel with the best quality (defined by the highest number of stars of the hotel), the lowest price, the lowest distance to the city center and the highest number of beds. A set of possible hotels is listed as follows.

| | Quality ($f_1(x)$) | Price ($f_2(x)$) | Centre Distance ($f_3(x)$) | Beds ($f_4(x)$) |
|---|---|---|---|---|
| Hotel Mozart ($x_1$) | 3 Stars | 110EURO/night | 6km | 400 beds |
| Hotel Verdi ($x_2$) | 2 Stars | 120EURO/night | 2km | 250 beds |
| Hotel Chopin ($x_3$) | 1 Star | 70EURO/night | 4.5km | 600 beds |
| Hotel Vivaldi ($x_4$) | 2 Stars | 105EURO/night | 1km | 200 beds |
| Hotel Beethoven ($x_5$) | 3 Stars | 124EURO/night | 6km | 350 beds |
| Hotel Bach ($x_6$) | 1 Star | 71EURO/night | 3km | 600 beds |
| Hotel Wagner ($x_7$) | 5 Stars | 170EURO/night | 3km | 100 beds |
| Hotel Hendel ($x_8$) | 3 Stars | 124EURO/night | 6km | 350 beds |
| Hotel Brahms ($x_9$) | 3 Stars | 124EURO/night | 6km | 250 beds |
| Hotel Schubert ($x_{10}$) | 2 Stars | 120EURO/night | 1km | 500 beds |

- Using Pareto-dominance, which of these hotels are non-dominated?
- If we normalize each objective, so all the values lie in the interval $[0, 1]$, what solutions will be non-dominated?

**Assignment 4 Distance Minimization Problem (DMP)**

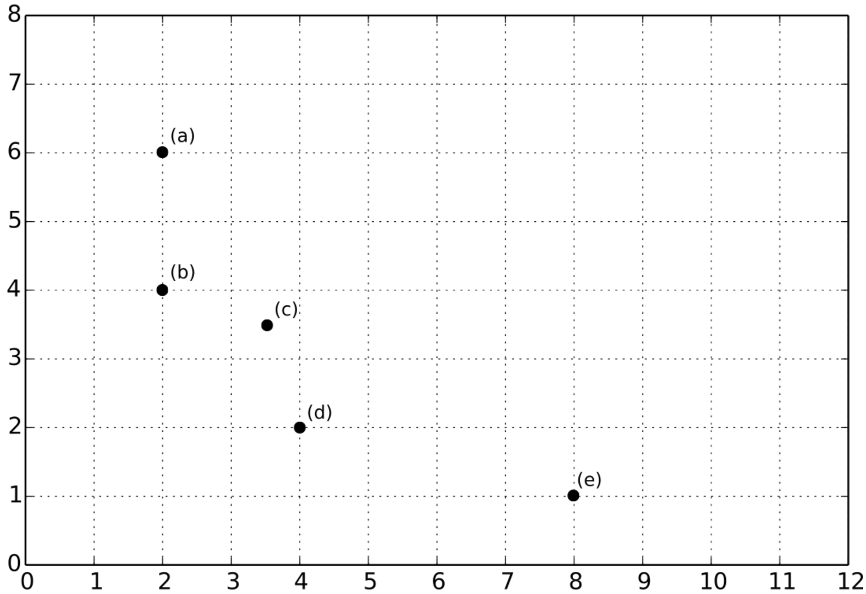
Assume you just moved to a new town and want to find a place to live for the next 3 years. However, you care about being close to certain locations in town in order to get the best living quality. It is very important for you to live close to the university and the supermarket.

- In order to be able to make a good choice for your new flat's location, you want to formulate this problem as a multi-objective optimization problem (explain which are the decision variables and the objective variables). Please describe how this can be done and explain where the optimal locations lie on a small example.
- In addition, you also want to be close to the gym. With these three points, what would be the new multi-objective optimization problem? Please explain how the Pareto-optimal solutions change compared to the previous case with only two points of interest.

**Assignment 5 A-Priori Methods**

In the lecture we have learned about different methods to solve optimization problems. In the following, we pay attention to a priori methods.

- Explain the difference between a priori methods, interactive methods and a posteriori methods.
- Explain how the Weighting Method from the lecture works. Which advantages and disadvantages does this have? In the following set of solutions, where both objectives are minimized, which solution(s) are optimal if we use the Weighting methods with the weights $(w_1, w_2) := (0.5, 0.5)$? Which ones are optimal for the weights $(0.1, 0.9)$?




- Explain how the $\varepsilon$-Constraint Method from the lecture works. Which advantages and disadvantages does this have? Which solution(s) are optimal if we optimize $f_1$ and set a constraint for $f_2$ of $\varepsilon_2 := 5.0$? How does this change if we change the constraint to $\varepsilon_2 := 3.0$?
- Explain how the lexicographic ordering of objectives from the lecture works. Which advantages and disadvantages does this have? In the above example, wich solution(s) are optimal if we specify the importance of objectives as $f_1 >> f_2$? Which solution(s) are optimal if we reverse the order?

**Assignment 6 A Priori Methods in Higher Dimensions**

Looking back to our holiday trip to Munich (Assighment 3), we would now like to apply some a priori method to help us pick a good hotel with the best quality (defined by the highest number of stars of the hotel), the lowest price, the lowest distance to the city centre and the highest number of beds. For this, we have normalized the features of each hotel (using min-max normalization). The set of possible hotels is listed as follows.

| | $f_1$ Normalized Quality (to be maximized) | $f_2$ Normalized Price (to be minimized) | $f_3$ Normalized Distance (to be minimized) | $f_4$ Normalized Beds (to be maximized) |
|---|---|---|---|---|
| $x_1$ | 0.5 | 0.4 | 1 | 0.6 |
| $x_2$ | 0.25 | 0.5 | 0.2 | 0.3 |
| $x_3$ | 0 | 0 | 0.7 | 1 |
| $x_4$ | 0.25 | 0.35 | 0 | 0.2 |
| $x_5$ | 0.5 | 0.54 | 1 | 0.5 |
| $x_6$ | 0 | 0.01 | 0.4 | 1 |
| $x_7$ | 1 | 1 | 0.4 | 0 |
| $x_8$ | 1 | 1 | 0.4 | 0.5 |
| $x_9$ | 1 | 1 | 0.4 | 0.3 |
| $x_{10}$ | 1 | 1 | 0.4 | 0.8 |

- Using the a priori weighting method on the normalized problem, with a weight of 0.2 for quality, a weight of 0.5 for price, a weight of 0.2 for distance to the city centre and a weight of 0.1 for number of beds; which of the hotels are optimal?
- Using the a priori $\epsilon$-constraint method, which hotels are optimal if we set $\epsilon_1 = 0.5$ and $\epsilon_3 = 0.5$ ?

## Assignment 7 Cone-Domination

Revise the concept of cone domination from the lecture and answer the following questions.

- What is the concept of cone domination and how does it work?
- Using the solutions shown in Assignment 5, which of the points are non-cone-dominated using the following matrix
$$\begin{pmatrix} 1 & 0.3 \\ 0.3 & 1 \end{pmatrix}$$
How does this compare to normal Pareto-dominance?
- Using the solutions shown in Assignment 5, which solution are non-cone-dominated with the following matrix
$$\begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$
- Which matrix has to be used in cone domination if we want to dominate solutions with a symmetric angle of $\varphi = 160$ degrees.

## Assignment 8 $\epsilon$-Domination

Revise the concept of $\epsilon$-domination from the lecture and answer the following questions.

- What is $\epsilon$-domination and how does it work?
- Using the solutions shown in Assignment 5, which of the points are non-$\epsilon$-dominated using $\epsilon = 0.5$?
  Note: Use the first definition given, not the simplified version.
- Using the solutions shown in Assignment 5, which of the points are non-$\epsilon$-dominated using $\epsilon = 1$? What is the smallest $\epsilon$-approximated Pareto front in this case?
  Note: An $\epsilon$-approximated Pareto front $PF_\epsilon$ consists of a set of solutions where for any solution $u$ in the problem, there exists a solution $v \in PF_\epsilon$ such that $v$ $\epsilon$-dominates $u$.
- What $\epsilon$ value can be used in order to assure that only solutions b and d are non-$\epsilon$-dominated?

## Assignment 9 Problem Formulation

Assume you work in a car manufacturing company, and your task is to design the structure of a new car. Your goal is to minimize the wind resistance of the car, while at the same time using only a fixed amount of material. The 3-D model of the car can be described by 7 positive real-valued parameters. Because there is the restriction of using at most a fixed amount of material, the sum of these parameters cannot be higher than 100. To test the wind resistance of a concrete assignment of these parameters, the values first have to be transformed into a 3-D model of a car. Then, using a simulation environment, the wind resistance can be determined.

- You have decided to use an Evolutionary Algorithm to solve this optimization problem. Please indicate which parts in the above description correspond to (1) the search space $S$, (2) the feasible search space $F$, (3) the decoded search space $W$, (4) the solution space $G$, (5) the decoding function $d$ and (6) the cost function $g$ of this problem.
- Describe a suitable representation of solutions and the optimization problem mathematically. What is the fitness function?
- What could be a possible neighborhood function that does not leave the feasible space $F$ when starting from a solution in $F$.

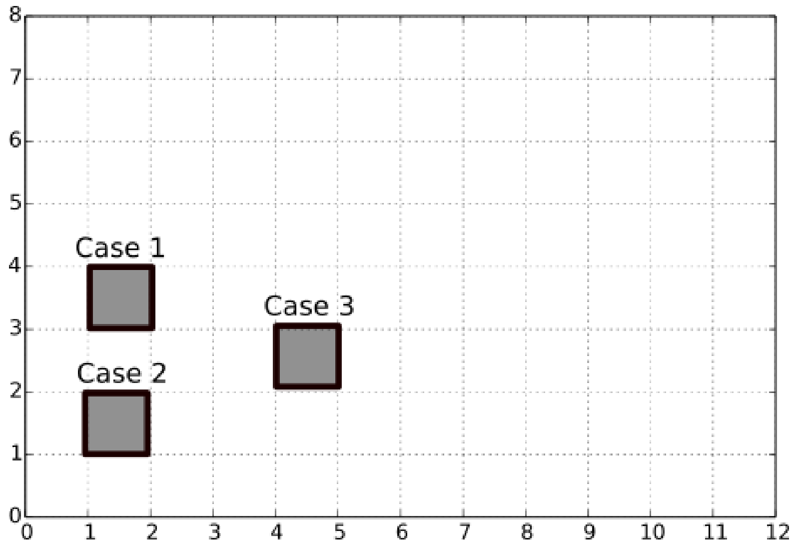
**Assignment 10        Interactive Methods**

Review the information about interactive methods and answer the following questions.

- Present three possible forms of interaction with the decision maker (DM).
- What characteristics should a good interactive method possess regarding the interaction with the DM?
- We want to solve the following bi-objective optimization problem:

$$\min(x_1 + 2x_2 + 3x_3, -2x_1 - 3x_3 + 6)$$
$$s.t. x_1, x_2, x_3 \in \{0, 1\}$$

  For which we use a binary encoding with three bits, and a standard deterministic 1-bit neighborhood. In a first iteration, the solutions $\vec{a} = (1, 0, 0)$, $\vec{b} = (0, 1, 0)$, $\vec{c} = (0, 0, 1)$, $\vec{d} = (0, 1, 1)$ and $\vec{e} = (1, 0, 1)$ were obtained.
  Then, only the two most relevant between these solutions should be presented to the DM. For deciding which ones, the DM provides a rectangular region in the objective space that he/she considers of higher interest. What solutions would you present in each of the following cases?




- For the same problem, each of the five solutions are updated by taking a solution in their neighborhood or keeping the current solution. This is done based on the region provided by the DM, so the selected solution is the best possible according to the DM preferences. Which new five solutions would you obtain for each of the previous cases?
  Note: When selecting the best solution in the neighborhood, a non-dominated solution (in the neighborhood) is always preferred, and the closeness to the region is used as a second factor.

**Assignment 11 Hill-Climbing**

- What are the advantages and disadvantages of using the Hill Climbing algorithm as an optimization algorithm?
- Apply the Hill Climbing algorithm to the TSP. The distance matrix $mat_{dist}$ indicates the distances between the cities $c_1$ to $c_5$:

$$\mathrm{mat}_{dist} = \begin{bmatrix} 0 & 110 & 350 & 220 & 70 \\ 110 & 0 & 455 & 260 & 170 \\ 350 & 455 & 0 & 420 & 490 \\ 220 & 260 & 420 & 0 & 170 \\ 70 & 170 & 490 & 170 & 0 \end{bmatrix}$$

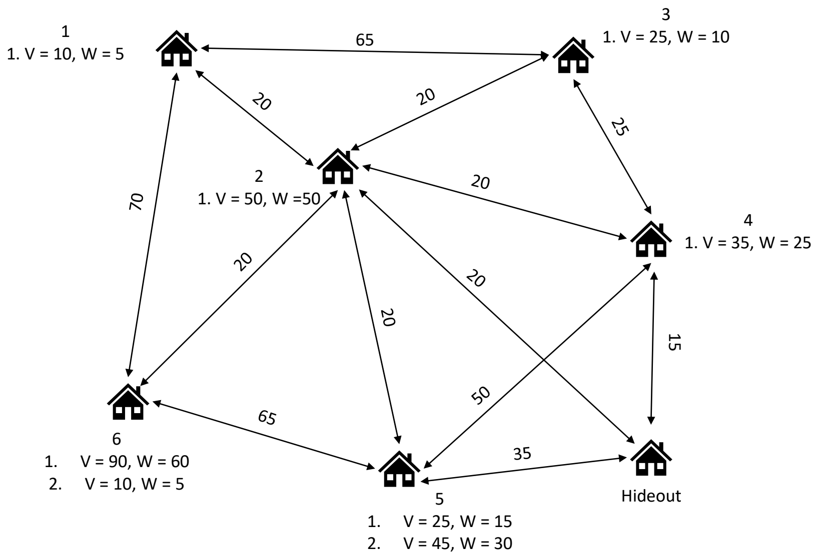
  where $c_1$: Los Angeles, $c_2$: San Diego, $c_3$: San Francisco, $c_4$: Las Vegas, $c_5$: California City.

  The Salesman does not need to come back to the city where he started, i.e. after visiting all five cities he will stay in the last city.

  Use the permutation representation of the TSP and exchange two cities as a neighborhood function. The fitness function of the problem is the sum of all distances, which should be minimized. Start with $c_2$-$c_4$-$c_1$-$c_5$-$c_3$ as initial solution $x_c$ and apply the algorithm until the stopping criterion is fulfilled.

**Assignment 12 Traveling Thief Problem**

Assume you are part of the most famous thieves guild in Germany. You and your fellow thieves have come up with a plan to rob several locations in one night and retire. Each location has some valuable items. You know the value of the items along with their weight in kilograms. You have created a graph were the nodes correspond to the locations and the edges represent the distances between locations:




Everyone in the guild was assigned a task and you are responsible for planing the operation, while keeping certain things in mind:

- You must start from the Hideout and return to the Hideout after the operation.
- The team executing the operation can only carry a maximum weight of $Q = 200[kg]$
- The team start out with a maximum speed $V_{max} = 1$
- The team speed is decreased as the weight they carry increases $V = V_{max} - q/Q$ ($q \equiv$ the current weight the team is carrying)
- The time taken to go from one location to another is $t = V * distance$

Your goal is to maximize your profit and minimize the time it takes to execute the operation, so the police has a lower chance of catching you. How can you model and formulate this problem? Which encoding and neighborhood function would you choose? Explain your reasoning!

## Assignment 13 Basics of Evolutionary Algorithms

- Considering the evolutionary optimization vocabulary, please indicate what are the genotype, phenotype, individuals, genes and alleles, and their relation with the problem formulation explained in the previous assignment sheet.
- Describe the general concept of how an EA works. Which components are necessary for an EA, and how do they work together? Visualize your concept.
- In an EA, selection mechanisms are used in two different parts of the algorithm. Explain what is the purpose of these two selections and what are the differences.
- Would the EA still reach its goal if you replaced one or both of these selection methods by a purely random selection? Justify your answer.

## Assignment 14 Selection Mechanisms and Selection Pressure

Consider the following (single-objective) optimization problem, which we want to optimize using an Evolutionary Algorithm (EA):

$$\begin{aligned} max \quad & f(x) = -x^4 + 8x^2 + 10x + 1000 \\ & x \in [-8, 8] \end{aligned}$$

This function has a local maximum at $x \approx -1.5425$ and a global maximum around $x \approx 2.2597$. The population of an EA in a given time step $t$ consists of the following eight solutions:

$$\begin{array}{llll} x^{(1)} = -7.0 & x^{(2)} = -2.0 & x^{(3)} = -1.0 & x^{(4)} = -0.5 \\ x^{(5)} = 1.3 & x^{(6)} = 2.0 & x^{(7)} = 2.5 & x^{(8)} = 3.0 \end{array}$$

- Calculate the fitness values and the relative fitness values of this population.
- Is the Roulette Wheel selection applicable to this problem? Which modifications might be necessary?
- Select from this population three individuals with Roulette Wheel selection and three individuals with Tournament selection. Explain how these two methods work.
- Take a look at the concept of selection pressure from the lecture. Which differences do you expect for the convergence speed (= number of generations the EA might need to reach the global maximum) if Roulette Wheel selection or Tournament selection is used in this situation?

## Assignment 15 Other selection mechanisms

For a **maximization problem** we have the following population of an evolutionary algorithm, which consists of 10 individuals (A to J):

| Indiv. | Fitness | Indiv. | Fitness |
|---|---|---|---|
| A | 5 | F | 4 |
| B | 13 | G | 8 |
| C | 3 | H | 20 |
| D | 2 | I | 1 |
| E | 4 | J | 3 |

Please answer the following questions:

- Explain the Universal Stochastic Sampling method and which advantage does it have regarding Roulette Wheel selection.
- We want to select 7 individuals from this population using Stochastic Universal Sampling with the given order of the individuals (A-J). The initial random number that is needed for this approach is chosen as $v = 3.0$. Which 7 individuals are selected?
- If the individuals in the first column (A-E) are the parents, and the individuals in the second column (F-J) are the offspring, what solutions will be selected for the next generation with $(\mu, \lambda)$-selection? Which ones will be selected with $(\mu + \lambda)$-selection?

## Assignment 16 Crossover Methods

An important part of an EA is the recombination. In the lecture, you have seen a number of different crossover operators for this purpose. Please answer the following questions.

- What is the purpose of a crossover operator in an EA?
- What is the difference between a crossover operator in a single-objective and in a multi-objective problem? Which modifications to the crossover might be necessary in multiple objectives?
- You are working in a company that specializes in the logistics area and your current task is to design an EA to solve a Traveling Salesman Problem (TSP). It was already decided that the permutation representation is used for this problem, as shown in the lecture. Now you need to decide for a suitable crossover method. Your boss asks you to use a one point crossover. What is your reaction? Please explain why this choice might be good or bad.
- You are optimizing a problem where a solution is represented through a 2-dimensional real-valued vector. For the two individuals $\vec{x} = (1.2, 5.0)$ and $\vec{y} = (0.4, 2.2)$, please perform
  - an offspring solution through arithmetic crossover with $\vec{\lambda} = (0.5, 0.8)$ and $\epsilon = 0.0$.
  - two offsprings using a one point crossover.

  Graphically show the locations of the produced solutions in the 2-dimensional decision space.

**Assignment 17 Genetic Algorithms**

In this assignment, you are tasked with calculating one whole generation of a genetic algorithm on the Traveling Salesman problem from the lecture. For the problem itself, you must start and end at the same city (first city in the sequence). Any city can be reached from any other city. The distances to the cities are given in the following distance matrix:

$$\mathrm{d} = \begin{bmatrix} 0 & 110 & 350 & 220 & 70 \\ 110 & 0 & 455 & 260 & 170 \\ 350 & 455 & 0 & 420 & 490 \\ 220 & 260 & 420 & 0 & 170 \\ 70 & 170 & 490 & 170 & 0 \end{bmatrix}$$

The fitness function is that of the traveling salesman problem, i.e. find a permutation of cities $\pi$ which minimizes the total tour length. Some additionally parameters for the algorithm:

- The population size is 5
- The initial population is :
  a) $c_1 - c_2 - c_3 - c_4 - c_5$
  b) $c_1 - c_3 - c_5 - c_2 - c_4$
  c) $c_1 - c_5 - c_2 - c_3 - c_4$
  d) $c_5 - c_1 - c_3 - c_2 - c_4$
  e) $c_2 - c_5 - c_4 - c_3 - c_1$
- Tournament selection is used as parent selection with a tournament size of $q = 2$
- For crossover, use Partially mapped crossover (PMX), with the partial sequence consisting of the third and fourth city in the sequence.
- For mutation, use swap mutation with a probability of $P_m = 0.1$
- Use a $(5 + 3)$-selection as reproduction scheme
- The stopping criteria checks if the best fitness has been updated in the last 5 generations.

Due to the random nature associated with this assignment, include a detailed step-by-step guide and the random numbers used for the various random elements. Also keep in mind that you should use this to present your solution.

## Assignment 18 Evolution Strategies

We will now take a look at Evolution Strategies. Make yourself familiar with the algorithm and answer the following questions.

- Explain how evolution strategies work. How are new solutions generated? How is the direction of exploration defined?
- The following real-valued function is to be minimized using the covariance matrix adaptation evolution strategy (CMA-ES) algorithm:

$$f(x_1, x_2) = (x_1 - 4)^2 + (x_2 - 2)^2$$

  The algorithms considers:

  - $\lambda = 4$ and $\mu = 2$,
  - the initial covariance matrix being the two-dimensional identity matrix,
  - the initial mean vector $m_0 = (0, 3)$,
  - the step size $\sigma = 1$,
  - a learning rate of 0.2,
  - the solution weights $w_1 = 0.7$ and $w_2 = 0.3$

  Do not use a normal distribution to sample the offspring of the initial $m$. Instead, produce four children from the initial solution by going step size into the positive or negative direction of the $x_1$ or $x_2$ axis (to take out the randomness factor of the algorithm and make everybodyÂ´s solutions comparable).

  What is the value of the updated $m$ and what does the covariance matrix look like after one iteration of the algorithm? What are the directions of the ellipsoid axes? (Hint: Calculate the Eigenvectors of the Covariance Matrix to get information about the directions.)

## Assignment 19 Particle Swarm Optimization

We are using a Particle Swarm Optimization (PSO) method to solve a minimization problem. The algorithm started at a time $t_0$ and is currently in time step $t_3$. The population consists of three particles $\vec{x}_1$, $\vec{x}_2$ and $\vec{x}_3$. In the following table you find the positions, velocities and fitness values of the particles at the time steps $t_0$, $t_1$, $t_2$ and $t_3$.

| Particle | $\vec{x}_1$ | $\vec{x}_2$ | $\vec{x}_3$ |
|---|---|---|---|
| $\vec{v}_i(t_0)$ | $(0.00, 0.00)$ | $(0.00, 0.00)$ | $(0.00, 0.00)$ |
| $\vec{x}_i(t_0)$ | $(2.00, 2.00)$ | $(-5.00, -2.00)$ | $(3.00, -5.00)$ |
| $f(\vec{x}_i, t_0)$ | 8.00 | 29.00 | 34.00 |
| $\vec{v}_i(t_1)$ | $(0.00, 0.00)$ | $(2.10, 1.20)$ | $(-0.30, 2.10)$ |
| $\vec{x}_i(t_1)$ | $(2.00, 2.00)$ | $(-2.90, -0.80)$ | $(2.70, -2.90)$ |
| $f(\vec{x}_i, t_1)$ | 8.00 | 9.05 | 15.70 |
| $\vec{v}_i(t_2)$ | $(0.00, 0.00)$ | $(2.31, 1.32)$ | $(-0.33, 2.31)$ |
| $\vec{x}_i(t_2)$ | $(2.00, 2.00)$ | $(-0.59, 0.52)$ | $(2.37, -0.59)$ |
| $f(\vec{x}_i, t_2)$ | 8.00 | 0.62 | 5.96 |
| $\vec{v}_i(t_3)$ | $(-0.78, -0.44)$ | $(0.92, 0.53)$ | $(-1.02, 1.26)$ |
| $\vec{x}_i(t_3)$ | $(1.22, 1.56)$ | $(0.33, 1.05)$ | $(1.35, 0.67)$ |
| $f(\vec{x}_i, t_3)$ | 3.92 | 1.21 | 2.27 |

- Explain briefly how PSO works.
- For each particle $i$, determine in $t_3$ its previous best position $\vec{P}_i(t_3)$ and the previous best position in its neighborhood $\vec{P}_g(t_3)$. The PSO uses a fully connected neighborhood topology.
- Calculate the updated velocities ($\vec{v}_i(t_4)$) and positions ($\vec{x}_i(t_4)$) for the next iteration of the PSO. Use $w = 0.4$, $\phi_1 = 0.3$, $\phi_2 = 0.3$ and $c_1 = c_2 = 1$.

**Assignment 20 SPEA2**

Take a look at the concept and components of Strength Pareto Evolutionary Algorithm 2 (SPEA2) and answer the following questions.

- Describe how SPEA2 works. What additional components are necessary to convert the general EA into SPEA2?
- For a 2-objective minimization problem, during the execution of SPEA2 and after the evaluation of solutions, the individuals in the population and archive are the following:

| Solution | $f_1(x)$ | $f_2(x)$ |
|---|---|---|
| Population | | |
| $x_1$ | 2 | 6 |
| $x_2$ | 4 | 3 |
| $x_3$ | 5 | 2 |
| Archive | | |
| $a_1$ | 1 | 8 |
| $a_2$ | 6 | 5 |

Given that the archive size is $N_A = 3$, and using a $k = 2$ for the k-th nearest neighbor method, compute the fitness value for each individual and update the Archive.

## Assignment 21 $\epsilon$-MOEA

Make yourself familiar with the $\epsilon$-Multi Objective Evolutionary Algorithm and answer the following questions:

- How are mating solutions selected in $\epsilon$-MOEA?
- How many children are produced in one iteration?
- How are the update mechanisms performed on the population and the archive?
- For a two-objective problem where $f_1$ is to be minimized and $f_2$ is to be maximized, the following solutions compose the population and archive at time step 10:

| Solution | $f_1(x)$ | $f_2(x)$ |
|---|---|---|
| Population | | |
| $x_1$ | 3 | 5 |
| $x_2$ | 1 | 6 |
| $x_3$ | 2 | 7 |
| Archive | | |
| $a_1$ | 2 | 5 |
| $a_2$ | 3.5 | 10 |

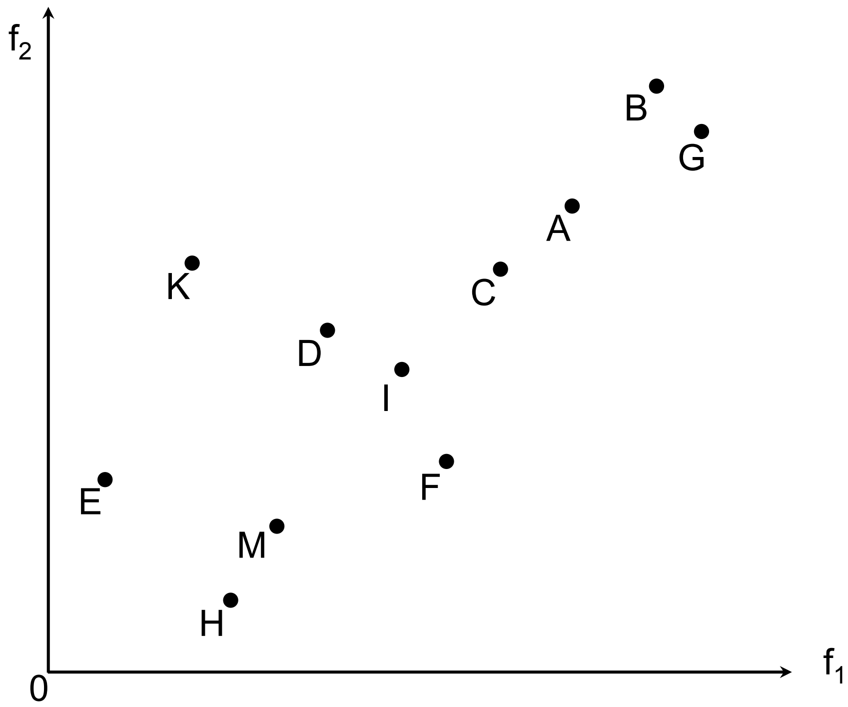
  The $\epsilon$-MOEA method being applied has an archive size of $N_A = 2$ and $\epsilon_1 = \epsilon_2 = 1$. After selecting solutions from the population and the archive, the crossover operation produces a child $x_{c10}$ at time step 10. The result of the evaluation of $x_{c10}$ is $f_1(x_{c10}) = 2.25$ and $f_2(x_{c10}) = 7.2$. Perform the population and archive update for time step 10 with the child $x_{c10}$. What are the solutions in the updated population and archive?

- At time step 11, a child $x_{c11}$ is generated, whose fitness values are $f_1(x_{c11}) = 2.5$ and $f_2(x_{c11}) = 8$. What are the solutions in the updated population and archive after time step 11? This task requires the correctly updated population and archive in time step 10.

### Assignment 22 NSGA-II

Take a look at the NSGA-II algorithm from the lecture and answer the following questions.

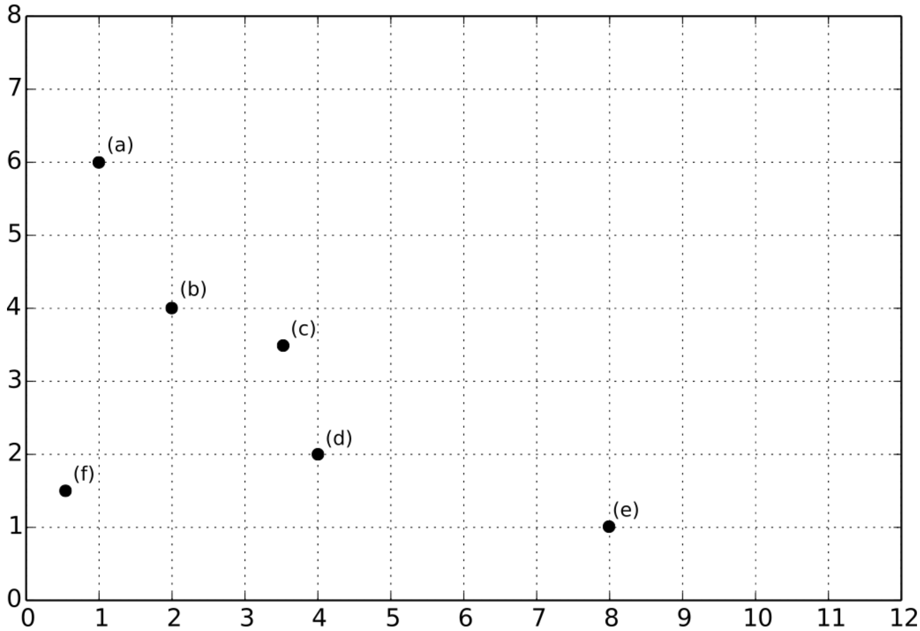
- Explain the basic concept of NSGA-II and how it ranks the solutions during the selection process.
- Given the following set of solutions in a multi-objective optimization problem, where both objectives should **be minimized**. Identify all the non-dominated fronts by using non-dominated sorting.



## Assignment 23 Crowding Distance

In NSGA-II, the concept of Crowding Distance is used. Please answer the following questions.

- Explain the purpose of this operator. What is it used for in NSGA-II and how does it work?
- Given the following set of solutions in a multi-objective optimization problem, where both objectives should **be minimized**. If NSGA-II is used with a population size of 5, which of the solutions will survive to the next generation?



## Assignment 24 Comparing MOEAs with PlatEMO

Make yourself familiar with both PlatEmo and the NSGA-II, SPEA2, $\epsilon$-MOEA, MOPSO and MOEA/D algorithms from the lectures. Using PlatEMO's built in GUI, run these algorithms on the distance minimization problem (MPDMP) from Assignment 4. Analyze the resulting pareto fronts and pareto sets graphically and try to explain the reasons for the results comparing the various algorithms. What can you conclude from the comparison?

Hints:

- To calculate the objective values, Euclidean distance is used.
- Set the number of objectives (M) to 2. After seeing what happens in 2 dimensions, change them to 3 and observe the changes in 3 dimensions.
- By default the fixed points for the 2-objective problem are: $ref1(-1, 0)$ and $ref2(1, 0)$.
- By default the fixed points for the 3-objective problem are: $ref1(0, 1)$, $ref2(0.8660, -0.5)$ and $ref3(-0.8660, -0.5)$.
- For your presentation, please show the fronts obtained with PlatEMO and share your opinion and analysis while presenting. As the analysis is done graphically, your submitted solution should show pictures of the fronts alongside some notes and observations.
- This is a more flexible assignment which should test your analytical skills.

**Assignment 25 MOEA/D**

Take a look at the Multi-objective Evolutionary Algorithm based on Decomposition (MOEA/D) from the lecture and answer the following questions.

- Please explain the basic concept of MOEA/D and how it tries to solve a multi-objective problem.
- What are the possible advantages and disadvantages of this approach?

In the following, MOEA/D is used for a **minimization** problem with 2 objectives, a population size of $N = 5$, a neighborhood size of $T = 3$ and an archive size of 5. The neighborhood $B(i)$ of each problem $i$ consists of its closest weight vectors in terms of Euclidean distance. The following table shows the weight vectors and the current population of MOEA/D. The archive consists of the same 5 solutions.

| $i$ | $\vec{\lambda}^i$ | $\vec{f}(x^i)$ |
|---|---|---|
| 1 | $(0.9, 0.1)$ | $(1.0, 7.0)$ |
| 2 | $(0.7, 0.3)$ | $(2.0, 4.0)$ |
| 3 | $(0.5, 0.5)$ | $(3.0, 3.0)$ |
| 4 | $(0.3, 0.7)$ | $(4.0, 2.0)$ |
| 5 | $(0.1, 0.9)$ | $(8.0, 1.0)$ |

- Identify the neighborhoods $B(i)$ for $i = \{1, .., 5\}$.
- During the main loop of the algorithm, for $i = 1$ we obtain a new solution $y$ from the recombination step with objective function values $\vec{f}(y) = (1.0, 6.0)$. Perform the neighborhood update and the update of the archive with this solution.

**Assignment 26 MOPSO**

We are using a Multi-Objective Particle Swarm Optimization (MOPSO) method to solve a 2-objective minimization problem. The algorithm is currently in time step $t_2$. The population consists of four particles $\vec{x}_1$, $\vec{x}_2$, $\vec{x}_3$ and $\vec{x}_4$. In the following table you find the positions, velocities and fitness values of the particles at the time steps $t_0$, $t_1$ and $t_2$.

| Particle | $\vec{x}_1$ | $\vec{x}_2$ | $\vec{x}_3$ | $\vec{x}_4$ |
|---|---|---|---|---|
| $\vec{v}_i(t_0)$ | $(0.0, 0.0)$ | $(0.0, 0.0)$ | $(0.0, 0.0)$ | $(0.0, 0.0)$ |
| $\vec{x}_i(t_0)$ | $(2.0, 2.0)$ | $(-3.0, 2.0)$ | $(5.0, -1.0)$ | $(-1.0, -2.0)$ |
| $f(\vec{x}_i, t_0)$ | $(5.0, 5.0)$ | $(3.0, 5.0)$ | $(7.0, 8.0)$ | $(2.0, 7.0)$ |
| $\vec{v}_i(t_1)$ | $(-1.0, -1.0)$ | $(1.0, -1.0)$ | $(-2.0, 0.0)$ | $(0.0, 1.0)$ |
| $\vec{x}_i(t_1)$ | $(1.0, 1.0)$ | $(-2.0, 1.0)$ | $(3.0, -1.0)$ | $(-1.0, -1.0)$ |
| $f(\vec{x}_i, t_1)$ | $(6.0, 2.0)$ | $(4.0, 5.0)$ | $(5.0, 2.0)$ | $(1.0, 6.0)$ |
| $\vec{v}_i(t_2)$ | $(1.0, -1.0)$ | $(1.0, 2.0)$ | $(-1.0, -2.0)$ | $(-1.0, 0.0)$ |
| $\vec{x}_i(t_2)$ | $(2.0, 0.0)$ | $(-1.0, 3.0)$ | $(2.0, -3.0)$ | $(-2.0, -1.0)$ |
| $f(\vec{x}_i, t_2)$ | $(8.0, 1.0)$ | $(5.0, 4.0)$ | $(6.0, 3.0)$ | $(2.0, 4.0)$ |

The archive $A(t_1)$ at $t_1$ consists of the following solutions:

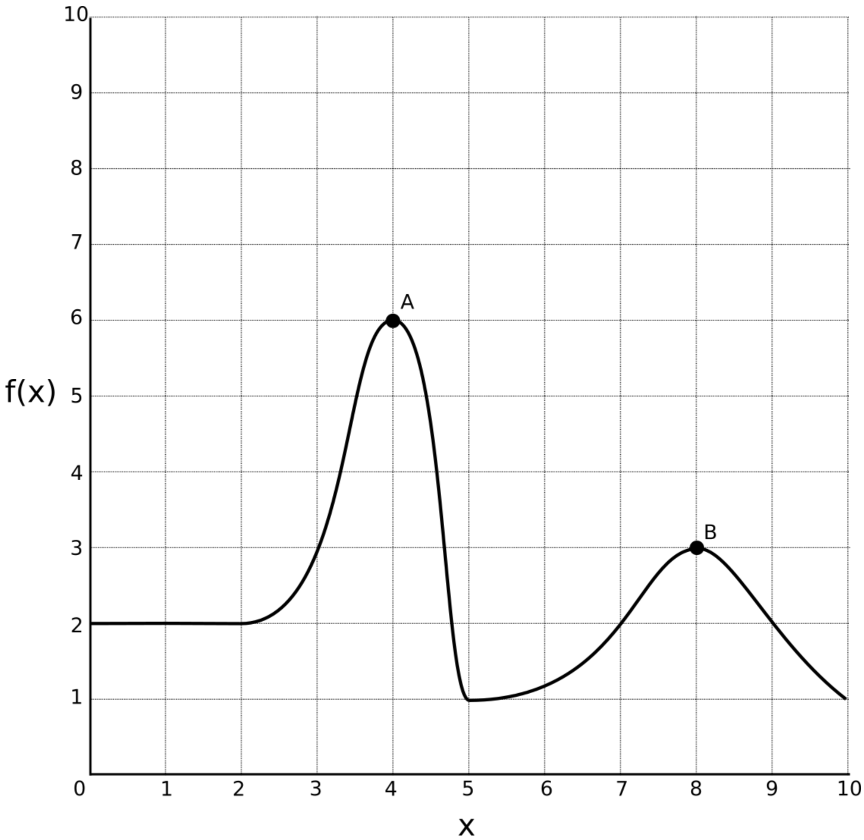
| $\vec{x}$ | $\vec{f}(\vec{x})$ |
|---|---|
| $(-1.0, -1.0)$ | $(1.0, 6.0)$ |
| $(-3.0, 2.0)$ | $(3.0, 5.0)$ |
| $(3.0, -1.0)$ | $(5.0, 2.0)$ |

- Update the archive $A(t_2)$ with the new solutions of time step $t_2$. The size of the archive is unlimited.
- The update of the personal best (cognitive component) is done by finding the best solution through a lexicographical ordering in which the first objective is the most important one, i.e. $f_1 \gg f_2$, for odd solutions ($\vec{x}_1$, $\vec{x}_3$); while even solutions ($\vec{x}_2$, $\vec{x}_4$) give more importance to the second objective. Determine the personal best $P_b(\vec{x}_2, t_2)$ for particle $\vec{x}_2$.
- The leader selection (social component) is done using the sigma method. The swarm uses a fully connected neighborhood. Select the leader $P_g(\vec{x}_2, t_2)$ for particle $\vec{x}_2$.
- Calculate the updated velocity ($\vec{v}_2(t_3)$) and position ($\vec{x}_2(t_3)$) for the next iteration of the MOPSO for particle $x_2$. Use $w = 0.5$, $\phi_1 = (1, 0.5)$, $\phi_2 = (0.5, 1)$, $c_1 = 0.3$ and $c_2 = 0.2$.

**Assignment 27 Robustness I**

In the lecture, we have learned about robustness of solutions in optimization. Please answer the following questions.

- What is robustness in the context of optimization? Why is robustness an important factor in single- and multi-objective optimization?
- What is the difference between type I and type II robustness from the lecture? Which additional element is necessary in an optimization algorithm to be able to use type II robustness?
- In the following, we show a plot of the fitness function in a single-objective maximization problem. For the two solutions $A$ (at $x_A = 4$) and $B$ (at $x_B = 8$), compute their respective mean effective fitness values $f^{eff}(x_A)$ and $f^{eff}(x_B)$. The neighborhood has a size of 3 and consists of the samples at $x + \delta$ for $\delta \in \{-1, 0, 1\}$.




- Which of the two solutions $A$ and $B$ is more robust according to type I robustness?
- Which of the two solutions $A$ and $B$ is more robust according to type II robustness, using a value of $\eta = 0.3$? Use $f^{eff}$ as the perturbed function.

## Assignment 28 Robustness II

Please answer the following questions related to Robustness.

- Explain the difference between robustness and reliability. Show a graphic example with at least one solution that is robust but not reliable, and another solution that is reliable but not robust.
- What is the difference between robustness in single- and in multi-objective problems? Explain what changes are needed in type I and type II robustness for multi-objective problems.
- In type I robustness, what is the influence of the parameter $\delta$?
- Discuss how the parameter $\eta$ influences type II robustness.

**Assignment 29 Robustness and Reliability in Multi-Objective Problems**

Consider the following bi-objective minimization problem with 2 decision variables and 1 inequality constraint:

$$\begin{aligned} min \quad & F(\vec{x}) = (f_1(\vec{x}), f_2(\vec{x})) = (|\sin(\frac{\pi}{2} \cdot x_1)|, \sqrt{x_1} + \frac{e^{x_2-1}}{3}) \\ subject\ to \quad & g_1(\vec{x}) = 8 \cdot x_1 + 2 \cdot x_2 - {x_1}^2 \geq 18 \end{aligned}$$

And the solutions $x_A = (2, 2)$ and $x_B = (5, 1)$. A Latin hypercube strategy is used for sampling, but instead of randomly selecting the solution inside each box, the upper right solution of the box is selected (e.g. for a box whose corners are $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$, the solution $(1, 1)$ is selected). The neighborhood has a size of $H = 4$ and $\delta_1 = \delta_2 = 1$. Please answer the following questions:

- What are the mean effective fitness values of $x_A$ and $x_B$? What can you say about the robustness of the solutions according to these mean effective fitness values?
- Using the same Latin hypercube strategy, what is the reliability value of $x_A$ and $x_B$?
- What should be the value of the required probability $R$ so both solutions $x_A$ and $x_B$ are reliable? What value $R$ would make both of them unreliable?

## Assignment 30 Constraint Handling

Take a look at the constraint handling techniques from the lecture and answer the following questions.

- What is constraint handling, and why is it needed in optimization algorithms?
- Consider the following single-objective problem with 2 decision variables and 2 inequality constraints.

$$\begin{aligned} \textit{min} \quad & f(\vec{x}) = 3 \cdot x_1 + 4 \cdot x_2 \\ \textit{subject to} \quad & g_1(\vec{x}) = 0.4 \cdot x_1 - 0.6 \cdot x_2 \geq 0 \\ & g_2(\vec{x}) = -0.1 \cdot x_1 + 0.8 \cdot x_2 - 0.1 \geq 0 \end{aligned}$$

For this problem, we decided to use the static penalty method with one fixed value of $R = 3$ as shown on slide EMO-6-35. For the three solutions $\vec{x}^{(a)} = (2, 4)$, $\vec{x}^{(b)} = (1, 5)$ and $\vec{x}^{(c)} = (7, 2)$, please compute the fitness of the solutions, their constraint violations for each of the two constraints and the penalized objective function values.

**Assignment 31 Constraint Handling in Multi-Objective Optimization**

Consider the following bi-objective minimization problem with 2 decision variables and 1 inequality constraint:

$$\begin{aligned} \textit{min} \quad & F(\vec{x}) = (f_1(\vec{x}), f_2(\vec{x})) = ({x_1}^2 + {x_2}^2, |x_1 - x_2 - 2 \cdot x_1 \cdot x_2 + 2|) \\ \textit{subject to} \quad & g_1(\vec{x}) = |x_1 + x_2| > 2 \end{aligned}$$

And a population consisting of the following eight solutions:

$$\begin{array}{llll} x^{(a)} = (0,0) & x^{(b)} = (1,1) & x^{(c)} = (2,2) & x^{(d)} = (-3,0) \\ x^{(e)} = (1,2) & x^{(f)} = (-1,0) & x^{(g)} = (3,-2) & x^{(h)} = (1,3) \end{array}$$

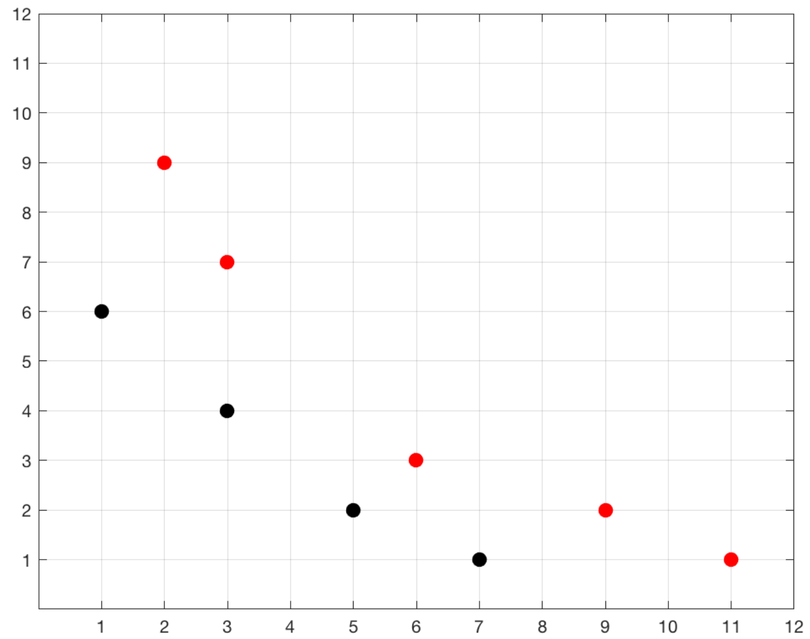
Take a look at the constraint handling techniques in MOPs and answer the following questions.

- What advantages and disadvantages do you find between using the penalty function for each objective and constrain-domination?
- Which solutions of the population are constrain-dominated by solution $x^{(d)}$?
- Which solutions are constrain-dominated by solution $x^{(b)}$?
- Which solutions are non-dominated according to constrain-domination?
- Which solutions are the most reliable?

## Assignment 32 GD and IGD

Take a look at the Inverted Generational Distance (IGD) and Generational Distance (GD) performance indicators from the lecture and answer the following questions.

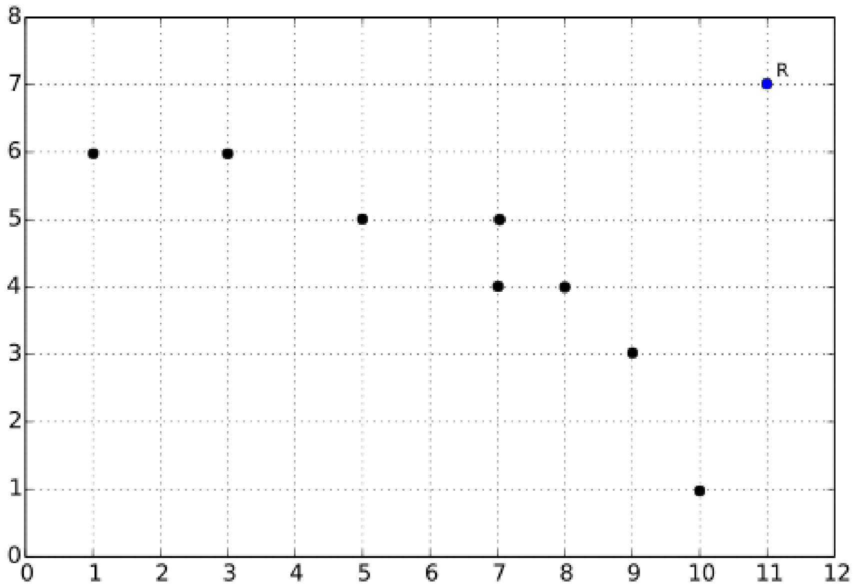
- Please explain how each of these metrics works, and what are the requirements for computing them.
- Please describe the difference between the IGD and the IGDX performance indicators.
- In the following, you see a set of solutions produced by an algorithm (red points) and a sample of the Pareto-optimal solutions (black points) in a multi-objective optimization problem. Both objectives should be minimized. Compute the GD and IGD values of these solution sets, using $q = 1$.



## Assignment 33 Hypervolume

Take a look at the Hypervolume (HV) metric from the lecture and answer the following questions.

- Please explain how this metric works, and what you need to compute it.
- Given the following set of solutions (black points) in a multi-objective optimization problem, where both objectives should **be minimized**, compute the HV of these points using the reference point R (in blue).




- Now try computing the HV of these points using the vector of worst objective function values (nadir point) as a reference point. What differences do you notice when using different reference points? Which of these two points is better for reflecting the actual quality of the solution set?
- What would happen if the reference point is too close to the solution set, i.e. the reference point is dominating the nadir point?

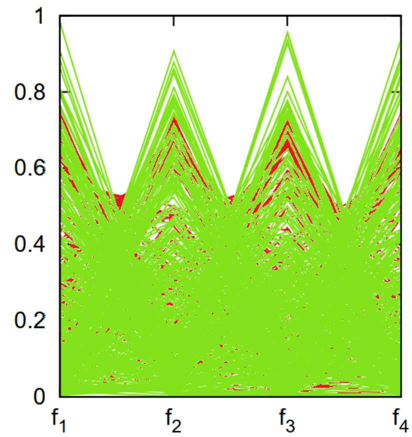
### Assignment 34 Experimental Evaluation

Read carefully the following text related to the experimental analysis of the results for a Multi-Objective Optimization problem:

> *We invented a new Multi-objective Evolutionary Non-dominated Sorting Algorithm (MENSA), which is intended to perform well with many objectives. We compare it against NSGA-II using the DTLZ2 benchmark with 4 objectives. The parameter setting of MENSA and NSGA-II is presented in the following table.*

| | MENSA | NSGA-II |
|---|---|---|
| Population size | 500 | 450 |
| Selection method | Tournament Selection | Tournament Selection |
| Crossover probability | 95 % | 90 % |
| Mutation probability | 5 % | 5 % |
| Max. number of generations | 150 | 100 |

> *With this configuration, we ran both algorithms a single time each, obtaining 500 solutions for MENSA and 450 solutions for NSGA-II. To visualize the obtained solutions, we use parallel coordinates, where MENSA is represented in green and NSGA-II in red.*




> *With this technique, we see that MENSA got better solutions than NSGA-II. Now, in order to measure the performance of MENSA, we used the GD metric with both algorithms, obtaining that GD(MENSA) $= 3.2$ and GD(NSGA-II) $= 2.7$. So it can be concluded that MENSA is better than NSGA-II.*

Please describe what parts of this experimental analysis were wrongly or vaguely performed. How could you improve this experiment?

## Assignment 35 Analyzing MOEAs Performance with PlatEMO

In this assignment, you will use the experiment module in PlatEMO for comparing the performance of MOEAs in different benchmarks. Your study must follow this steps:

- Compare the five algorithms from the lectures (NSGA-II, SPEA2, $\epsilon$-MOEA, MOEA/D and MOPSO). For MOPSO, MOEA/D $\epsilon$-MOEA use the default parameters. Use a population of 100 individuals, and 10000 evaluations.
- The comparison consists of three experiments:
  - In a first experiment, you will compare the benchmarks ZDT1, ZDT2, ZDT3, ZDT4 and ZDT6 using their default number of decision variables.
  - A second experiment will evaluate the benchmarks DTLZ1, DTLZ2, DTLZ3, DTLZ4 and DTLZ6 with 3 objective variables and the default number of decision variables.
  - The last experiment consists of evaluating again the benchmarks DTLZ1, DTLZ2, DTLZ3, DTLZ4 and DTLZ6, but with 10 objective variables and the default number of decision variables.
- You must compare the algorithms in terms of their GD, IGD and runtime.
- For each test, 30 runs of each algorithm must be performed, and the average, standard deviation and statistical significance results must be computed. For the statistical significance, you must select one algorithm to be compared with the rest. We recommend selecting the one that you consider the best (but this one has to be the same in all the experiments).
- The experiments may take a long time (especially the third one). So it is recommended to run in parallel and be patient.
- Finally, you will analyze the obtained results and draw some conclusion on which algorithms work better for different types of problems.