University of Exeter

FACULTY OF ENVIRONMENT, SCIENCE AND ECONOMY

COMM510

Multi-Objective Optimisation & Decision Making

Continuous Assessment

Date Set: 26th October 2023
Date Due: 13th December 2023

This CA comprises 40% of the overall module assessment.

This is an **individual** exercise. You are required to cite the work of others used in your solution and include a list of references, and must avoid plagiarism, collusion and any academic misconduct. Your attention is drawn to the ELE online module on Academic Honesty and Plagiarism, which all students must complete (`https://ele.exeter.ac.uk/course/view.php?id=1957`).

The assessment for COMM510 covers multiple aspects of research in multi-objective optimisation and decision making via a small research project. This assignment includes engaging with literature (independent reading); formulating a research programme addressing a particular question; reporting on computational experiments; analysis of results; relating results to existing literature; identifying future research directions given the results. The assignment is *summative*. Please ensure you read the entire document before you begin the assessment.

1 Assignment introduction

In this coursework you will undertake a small research project, addressing one of 10 topics posed in Section 3 of this document. This will entail the use of existing packages for optimisation and (to varying degrees) developing your own code. You are not restricted to particular programming languages/implementations – however, all topics involve the analysis and presentation of results from computational experiments and necessitate further reading exploring in-depth various aspects of multi-objective optimisation and decision making that have been touched upon in the lecture series. Please refer to the module specifier for further details on intended learning outcomes (ILOs).[1] All submissions **must** be anonymous (when detailing author information in the LaTeX document, please use your student number rather than name). You may of course ask further questions regarding topics with module staff during the weekly office hours.

2 Submission structure

This coursework requires the submission of a concise document including a background and literature review on a research topic you have chosen from the list in Section 3. The submission must also include a description of the research programme you have undertaken to address the research topic, including aspects such as detailing the broader research question you will examine, the experimental design, the presentation of the experiments, the analysis of the results, a discussion and any conclusions, and setting out future work directions given these results. The coursework marking criteria are at the end of this document.

The submission must adhere to the short paper style of the ACM GECCO conference, which is a double-column format with 4 pages maximum (excluding references). The ACM LaTeX style file will be made available on the COMM510 ELE page, which will be used in this coursework, along with an initial `.tex` file to work from.

Precisely, the document (maximum of four pages, excluding references) should include:

- A short abstract outlining the document.
- An introduction.
- A review of the literature about the topic, describing the background to the chosen research question and what work has been done in the existing literature to investigate it.
- A reasoned plan of the empirical work undertaken, including (as suitable), algorithms to compare, test problems to use, quality measures to employ, experimental protocols, how the results will be evaluated, etc.
- Your prior expectations of the results, given any insights from the literature and your understanding of the task and research question. This might take the form of a research hypothesis to be investigated.
- A presentation of the results obtained in an appropriate form (e.g. tables, plots, etc.).
- An analysis of the results.
- A contextualisation of the results, relating them to the existing literature.
- A conclusion which outlines potential future research directions that lead on from your work.

3 Topics

Below is the list of research topics for the COMM510 coursework. You must select one topic for this assignment.

1. **The impact of initialisation on MOEAs**. In this project topic you should explore the research question: how much of the computational budget (objective function evaluations) should be spent on initialisation? Typically the number of random initial solutions (if no existing solutions are usable), is the same as the search population – however it is possible to generate substantially more initial random solutions and take the "N" best of these to form the initial search population. You will want to explore the effect of varying the amount of budget you assign to initialisation versus search for one or more optimisers. You may also want to consider the process by which the initial designs are generated.

[1] `https://intranet.exeter.ac.uk/emps/studentinfo/subjects/computerscience/modules/2023/index.php/?moduleCode=COMM510`

2. **How amenable/exploitable are test problems suites to seeding?** In this project topic you should explore the research question: how much easier are test problems from a suite if search is initialised from a set of solutions which include a Pareto optimal decision vector, and investigate why the results are as observed. You should contrast at least two multi-objective test *suites* in this project.

3. $(1+1)$ **or** $(|A|+1)$**?** In this project topic you should explore the research question: how (and why) does the relative performance of the simple "greedy" multi-objective evolution strategy optimiser vary if a $(1+1)$ formulation or a $(|A|+1)$ formulation is used. In a $(1+1)$ (as in PAES), the child replaces the parent if it dominates it, or if is non-dominated with the approximation set but is in a preferred region (e.g. in a less densely populated region in objective space, contributes more to the hypervolume, etc.), whereas in the $(|A|+1)$ approach the parent is newly drawn each generation from the approximation set. You should consider a range of test problems, and also the effect of *how* you draw parents from the archive/approximation set A.

4. **Does correlation help in MOEAs?** This project aims to explore the effect of correlations on the performance of multi-objective evolutionary algorithms. Different problems can have different correlations among objective functions. For example, anti-correlated means they are conflicting and fully correlated means they are not conflicting. In this project, you will test different multi-objective evolutionary optimisers on MOPs with different correlations among objective functions. The expectation of the project is to provide some guidelines in selecting a MOEA based on the correlation among objective functions. You may not find problems with correlation values; therefore, you will need to estimate the correlation, e.g., by using Kendall's correlation measure. You can select three MOEAs (one in each category) and five MOPs. For example, you can select a range of test problems with different correlations among objective functions.

5. **Does correlation help in Multi-objective Bayesian Optimisation?** This project aims to explore the effect of correlations on the performance of multi-objective Bayesian Optimisation. Different problems can have different correlations among objective functions. For example, anti-correlated means they are conflicting and fully correlated means they are not conflicting. In this project, you will use multi-task Gaussian processes with an acquisition function of your choice (e.g., expected hypervolume improvement) and test the algorithm on MOPs with different correlations among objective functions. The expectation of the project is to compare independent Gaussian process models and correlated models (after using multi-task GPs) in multi-objective Bayesian optimisation. You may not find problems with correlation values; therefore, you will need to estimate the correlation, e.g., by using Kendall's correlation measure. You can select five MOPs to test the algorithm. For example, you can select a range of test problems with different correlations among objective functions.

6. **Mono- and multi-surrogate approaches for scalarising functions** This project aims to explore the performance of mono- and multi-surrogate approaches in multi-objective optimisation. Given a scalarising function, you will need to compare the performance of mono- and multi-surrogate approaches. This project specifically ask for three scalarising functions: Hypervolume improvement, dominance ranking and minimum signed distance. Check this paper `https://dl.acm.org/doi/pdf/10.1145/3071178.3071276` for more details. You can select five MOPs to compare the approaches. For example, you can select a range of test problems with a different number of objectives.

7. **Comparison of Bayesian models in single-objective Bayesian optimisation** This project aims to explore the performance of different Bayesian models in single-objective Bayesian optimisation. Gaussian processes as Bayesian models are widely used in Bayesian optimisation. However, their efficacy is limited to low-dimensional problems. Other Bayesian models such as Bayesian neural networks and Bayesian random forests can also quantify the uncertainty, which can be used in Bayesian optimisation. In this project, you will compare at least one Bayesian model in addition to the Gaussian process and three acquisition functions (e.g, expected improvement, probability of improvement, entropy search and lower confidence bound) in Bayesian optimisation. You can select five single-objective optimisation problems of your choice. For example, you can select a range of test problems with a different number of decision variables.

8. **Comparison of Bayesian models in multi-objective Bayesian optimisation** This project aims to explore the performance of different Bayesian models in multi-objective Bayesian optimisation. Gaussian

processes as Bayesian models are widely used in Bayesian optimisation. However, their efficacy is limited to low-dimensional problems. Other Bayesian models such as Bayesian neural networks and Bayesian random forests can also quantify the uncertainty, which can be used in Bayesian optimisation. In this project, you will compare at least one Bayesian model in addition to the Gaussian process with expected hypervolume improvement as the acquisition function in multi-objective Bayesian optimisation. You can select five multi-objective optimisation problems of your choice. For example, you can select a range of test problems with a different number of decision variables.

9. **Many-objectivisation in multi-objective Bayesian Optimisation** Exploitation and exploration are two main goals in both single and multi-objective Bayesian optimisation, which are simultaneously achieved by optimising an acquisition function (e.g., expected improvement and expected hypervolume improvement). Given a multi-objective optimisation problem (all objectives to be minimised), a different way to achieve exploitation and exploration is minimising posterior means (as exploitation) and maximising posterior standard deviations (as exploration) from Gaussian process models. It means a $k-$ objective optimisation problem will be converted to $2 \times k$ multi(or many)-objective optimisation problem. Use this approach of many-objectivisation in multi-objective Bayesian optimisation and compare it with the standard approach (e.g., by using EHVI). You can select five multi-objective optimisation problems of your choice. For example, you can select a range of test problems with a different number of objectives.

10. **Multi-objective Optimisation in Machine Learning** This task focuses on utilising multi-objective optimisation in Machine learning. Select a regression or classification algorithm of your choice and estimate its hyperparameters using a multi-objective optimiser. For example, architecture in neural networks including weights, number of layers and number of nodes, are the hyperparameters, which are usually estimated by optimising a single-objective function (also called loss function in machine learning). In this task, you will formulate a multi-objective optimisation problem on a machine learning method and solve it using an appropriate optimiser.

4 Software tools and packages

In planning your empirical work it is worth noting there are many existing open source packages containing implementations a number of pre-existing multi-objective optimisers, and test problems/suites.

5 Submission

The document body of the pdf format report should be no more than 4 pages (excluding references) in length. It should be typeset in LaTeX, using the style file provided on the COMM510 ELE page and be submitted by 12pm (midday) on the date specified on the cover page, using the ELE-based submission system.

If you are less familiar with LaTeX, we suggest you take advantage of the institutional Overleaf account `https://www.overleaf.com`, which you will benefit from if you register on the website with your university email address. This includes an online editor and compiler, version control, as well as 'how to' guides for typesetting using LaTeX (as covered in workshops).

Marking criteria are tabulated on the following page in Section 6.

6 Marking criteria

The assessment will be marked using the following criteria.

Abstract.	The degree to which the report and contents are concisely and clearly presented in the document abstract.	/5
Introduction and background literature review.	The degree to which the report introduction sets out the research question being set and the expectations to the reader of the rest of the documents, and within the background section an appropriate body of literature has been synthesised and presented, and the degree to which it contextualises and supports the research programme proposed, with no obvious omissions.	/20
Research Programme.	The degree to which the research programme presented is clear, well-aligned to the research question(s) being addressed and sufficient in scale/ambition, and has a clear and appropriate plan for the assessment of the empirical outcomes. The degree to which the expected outcomes are reasonable and well-justified given the published research in the domain.	/20
Results presentation.	The degree to which the presentation of results is clear and support the research topic being investigated, and support the reader alongside the analysis.	/20
Analysis and Contextualisation of results.	The degree to which the analysis provided of the results is appropriate, supported by the data, and provides insight into the research question, and the degree to which the results and analysis have been well-contextualised with existing results in the literature.	/20
Future research directions.	The degree to which sensible further research directions and questions have been mapped out, based upon the results and analysis provided in the report.	/5
Presentation.	The degree to which the report as a whole is clear, concise, well-written, well-structured, well-formatted and devoid of grammatical error.	/10
Length penalty.	A penalty of 10 marks will be applied for each page (or part thereof) that the document body (excluding references) is in excess of 4 pages.	
	Total	/100