Numerical Optimization for Large Scale Problems
Assignments

Politecnico di Torino, A.Y. 2025/2026

PLEASE, CAREFULLY READ ALL THE INSTRUCTIONS BEFORE PREPARING AND SUBMITTING YOUR REPORT

Recall that the assignment allows you to **score up to 12 points**, according to **the chosen assignment, the report contents, and the defense**. Suitably addressing all the points reported, allows you to score **up to** the corresponding points. Each team has to choose a single assignment either from the Unconstrained or the Constrained optimization section.

Contents

1 Guidelines

ATTENTION: The final score will also depend on the level of the report's compliance with the following guidelines.

1.1 General guidelines

1. The assignment can be done in **teams of at most 3 people**.
2. The final score of the assignment depends both on the report contents and the defense; then, **people belonging to the same team can obtain different scores**, according to their defense performance.
3. **People in the same team are expected to take the exam in the same call**. If a team member fails the exam or rejects the mark he/she can take the exam in another call maintaining the same report.

1.2 Guidelines for writing the report

1. The **document** is expected to report:
    1.1. a brief introduction to the report [**max. half a page**];
    1.2. a brief description of the implemented methods [**max. one page per method**];
    1.3. an introductory analysis of the test problems considered [**max. one page per test problem**];
    1.4. if implemented, a brief description of the finite differences implementation [**max. one page per test problem**];
    1.5. tables and/or figures summarizing your results [**see assignments for more details**]
    1.6. comments and analyses on your results for the test problems [**max. one page per test problem**].
2. You must write the **name, family name, and student ID of each team member** at the beginning of the report.
3. Please use **captions** in order to explain what every table and/or figure is reporting, and quote it also in the text (e.g., "In Figure xx we report the plot of....", "In Table yy we compare ...").
4. You are expected to **test your solvers** on some common problems, with different values of some parameters and possibly different starting points. In all the cases, you have to **compare the results obtained**, for example in terms of the number of iterations and computing time, commenting on your results also in view of the values of the parameters used and of the theory.
5. As an **appendix** of the report, **you must add the commented scripts/functions you implemented** in your favorite programming language. Please make sure to use sensible names for the variables and functions, and to provide enough comments and explanations to make **the code readable to a non-expert of the specific language**.

1.3 Submission guidelines

1. You are expected to submit **a single pdf file** per group (and **not** per person). **Avoid** compressed folders. If you upload something which is not a single pdf file, I'll ask you to resubmit.

2. The **name of the pdf file must be** "`NumOptReport2526_`" followed by the family names of the team members separated by underscores ("_"), in alphabetical order; for example, for a team of students with family names *Rossi*, *Bianchi*, and *Verdi*, the pdf file must be

   `NumOptReport2425_Bianchi_Rossi_Verdi.pdf`

   Do not use whitespaces in the file name.

3. The file should be submitted through the "**Consegna elaborati**" tab on the course page.

4. If you work in a group (**max 3 people**) please upload the file **only once** but clearly state in the file name the family names of all team-mates.

5. The **deadline** for submission is **one week before the date of the official call** at which you aim to take the exam.

1.4 Discussion guidelines

1. For the discussion of the report, **the team must bring a PC** with all the codes ready to be run, in case of needs, and the pdf of the uploaded report.

2. The team will have **12 minutes to present the results** contained in the report. After the presentation, there will be a discussion with the teacher, asking questions to the team members.

3. The team members will discuss the report directly commenting on it (i.e., on the complete pdf file), **without** supporting slides or other kind of summarizing presentations.

Bibliography

[1] https://www.researchgate.net/publication/325314497_Test_Problems_for_Unconstrained_Optimization

2 Assignment on Unconstrained Optimization

2.1 Assignment: Derivative-based Optimization (max 12 pts)

1. Implement **exactly two** out of the following numerical methods for unconstrained optimization:

   1.1. [**0.5 points**] Steepest Descent method (+ Back-tracking);

   1.2. [**1 point**] Nonlinear Conjugate Gradient method, either Fletcher-Reeves or Polak-Ribière, (+ Back-tracking);

   1.3. [**2 points**] Modified Newton method (+ Back-tracking);

   1.4. [**0.5 points**] Inexact Newton method (+ Back-tracking);

   1.5. [**2 points**] Truncated Newton method, a.k.a. line search Newton-Conjugate Gradient method (+ Back-tracking).

2. [**5.5 points**] Apply the codes to **exactly two test problems taken from [1]** following these instructions:

   - **The Rosenbrock function** or its variants in [1] are **not allowed**.
   - Computes the **exact derivatives** of the chosen functions and **report them in the introductory analysis of the test problems** (see Section 1.2).
   - **set a random seed** equal to the minimum student ID of the team members; e.g., given three team members with student IDs 9876, 1234, and 2345, the random seed must be set equal to $1234 = \min\{9876, 1234, 2345\}$.
   - study the problem for these **values of dimension**: $n = 2, 10^3, 10^4, 10^5$.
   - use $\boldsymbol{x}^{(0)} = \bar{\boldsymbol{x}} \in \mathbb{R}^n$ as **starting point**, where $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ is the starting point suggested in [1] for the problem. Then, use **5 new starting points randomly generated** with uniform distribution in a hyper-cube $[\bar{x}_1 - 1, \bar{x}_1 + 1] \times \cdots \times [\bar{x}_n - 1, \bar{x}_n + 1] \subset \mathbb{R}^n$.
   - Concerning any possible **parameter** of the methods implemented (e.g., Back-tracking parameters, forcing terms, stopping criteria, etc.), **tune it and discuss/motivate your choice**. If you perform preliminary tests with parameters that are not your final choice, you do not need to include the numerical results in the report, but you must briefly describe what happened in the preliminary tests.

3. Extend the study in the case where derivatives are computed by using **an efficient implementation of the finite differences** (general purpose finite differences are not recommended). **Report the strategy** you adopted for implementing the finite differences in a specific section/paragraph (see Section 1.2).

   In particular:

   - [**1 point - Only for methods with Hessian**] Use the exact gradient, but approximate the Hessian considering the following values for the increment h for each differentiation:

     $$h = 10^{-k}, \qquad k = 4, 8, 12\,,$$

     and also the case of the specific increment h_i when differentiating with respect to the variable x_i, according to these values:

     $$h_i = 10^{-k}|\hat{x}_i|, \qquad k = 4, 8, 12\,, \quad i = 1, \ldots, n\,,$$

     where $\hat{\boldsymbol{x}} = (\hat{x}_1, \ldots, \hat{x}_n) \in \mathbb{R}^n$ is the point at which the derivatives have to be approximated.

- [**1.5 points**] Approximate all the derivatives (i.e., also the Hessian if included in the method) considering the following values for the increment h for the gradient, for each differentiation:

$$h = 10^{-k}, \qquad k = 4, 8, 12\,,$$

  and also the case of the specific increment h_i when differentiating with respect to the variable x_i, according to these values:

$$h_i = 10^{-k}|\hat{x}_i|, \qquad k = 4, 8, 12\,, \quad i = 1, \ldots, n\,,$$

  where $\hat{\boldsymbol{x}} = (\hat{x}_1, \ldots, \hat{x}_n) \in \mathbb{R}^n$ is the point at which the derivatives have to be approximated.

4. After running the optimization methods according to the instructions above, realize a throughout comparison among the methods for each test problem (see item related to comments, Section 1.2), reporting data in tables and/or figures (see item related to tables/figures, Section 1.2). Comment on your results also in view of the expected theoretical behavior.

   For commenting the results, for each method, for each test problem, you can use **table templates** like Tables 1 and 2. You can add extra columns/information in the tables if you believe they/it can help the analysis.

   The following figures are **mandatory**:

   - top view of the function and of all the sequence paths, for the cases $n = 2$. Collect in the same figure the paths for each different starting point/simplex; create separate figures for the finite difference-based sequences. **Be sure that the figures are readable**.
   - Experimental rates of convergence for the sequences that converged. Collect in the same figure the rates different starting point; create separate figures for each dimension n, and for the finite difference-based sequences. **Be sure that the figures are readable**.

start.pt ID	**grad.norm**	**iters/max.iters**	**success**	**flag**	**rate of conv. (exp.)**	**time**
$\bar{\boldsymbol{x}}$	1.75e-7	57/1000	yes	-	1.75	25s
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Avg (successes)	1.5e-7	60/1000	-	-	1.5	30s

Table 1: Table template for collecting results - exact derivatives.

start.pt ID	k	**grad.norm**	**iters/max.iters**	**success**	**flag**	**rate of conv. (exp.)**	**time**
$\bar{\boldsymbol{x}}$	4	1.5e-5	173/1000	yes	-	1.1	40s
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\bar{\boldsymbol{x}}$	8	1.5e-5	166/1000	yes	-	1.1	38s
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\bar{\boldsymbol{x}}$	12	1.5e-5	195/1000	yes	-	1.1	35s
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Avg (successes)	4	1.25e-6	125/1000	-	-	1.02	55s
Avg (successes)	8	$\cdots$	$\cdots$	-	-	$\cdots$	$\cdots$
Avg (successes)	12	$\cdots$	$\cdots$	-	-	$\cdots$	$\cdots$

Table 2: Table template for collecting results - finite differences

2.2 Assignment: Nelder-Mead method (max 10 pts)

1. [**2 points**] Implement the Nelder-Mead method for unconstrained optimization.

2. [**8 points**] Apply the Nelder-Mead method to **exactly three test problems taken from [1]** following these instructions:

   - **The Rosenbrock function** or its variants in [1] are **not allowed**.
   - **set a random seed** equal to the minimum student ID of the team members; e.g., given three team members with student IDs 9876, 1234, and 2345, the random seed must be set equal to $1234 = \min\{9876, 1234, 2345\}$.
   - study the problem for these **values of dimension**: $n = 2, 10, 20, 50$.
   - use $\boldsymbol{x}^{(0)} = \bar{\boldsymbol{x}} \in \mathbb{R}^n$ as **starting point** for generating the starting simplex, where $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ is the starting point suggested in [1] for the problem. Then, use **5 new starting points randomly generated** with uniform distribution in a hyper-cube $[\bar{x}_1 - 1, \bar{x}_1 + 1] \times \cdots \times [\bar{x}_n - 1, \bar{x}_n + 1] \subset \mathbb{R}^n$.

     The starting simplices must have the starting points mentioned above as a vertex, while all the other vertices are randomly generated as $\bar{\boldsymbol{x}} + \eta_i \boldsymbol{e}_i$, for each $i = 1, \ldots, n$, where η_i is a random value in $[-1, -0.5] \cup [0.5, 1]$, obtained by sampling uniformly from $[0.5, 1]$ and randomly selecting the sign $\pm$.
   - Concerning any possible **parameter** of the Nelder-Mead method, **tune it and discuss/motivate your choice**. If you perform preliminary tests with parameters that are not your final choice, you do not need to include the numerical results in the report, but you must briefly describe what happened in the preliminary tests.

3. After running the optimization methods according to the instructions above, realize a throughout comparison among the test problems (see item related to comments, Section 1.2), reporting data in tables and/or figures (see item related to tables/figures, Section 1.2). Comment on your results also in view of the expected theoretical behavior.

   For commenting the results, for each test problem, you can use the **table template** of Table 1. You can add extra columns/information in the tables if you believe they/it can help the analysis.

   The following figures are **mandatory**:

   - top view of the function and of all the sequence paths, for the cases $n = 2$. Collect in the same figure the paths for each different starting point/simplex; create separate figures for the finite difference-based sequences. **Be sure that the figures are readable**.

3 Assignments on Constrained Optimization

3.1 Assignment: Projected gradient method (max 12 pts)

1. Implement **at least one** of the following methods:

   1.1. [**0.5 points**] Projected gradient with line search along the feasible direction;

   1.2. [**2 points**] Projected gradient with line search along the projection arc;

2. [**8.5 points**] Apply the codes to **exactly two test problems taken from [1]** following these instructions:

   - **The Rosenbrock function** or its variants in [1] are **not allowed**.
   - Computes the **exact gradients** of the chosen functions and **report them in the introductory analysis of the test problems** (see Section 1.2).
   - **set a random seed** equal to the minimum student ID of the team members; e.g., given three team members with student IDs 9876, 1234, and 2345, the random seed must be set equal to $1234 = \min\{9876, 1234, 2345\}$.
   - study the problem for these **values of dimension**: $n = 2, 10^3, 10^4, 10^5$.
   - use $\boldsymbol{x}^{(0)} = \bar{\boldsymbol{x}} \in \mathbb{R}^n$ as **starting point**, where $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ is the starting point suggested in [1] for the problem. Then, use **5 new starting points randomly generated** with uniform distribution in a hyper-cube $[\bar{x}_1 - 1, \bar{x}_1 + 1] \times \cdots \times [\bar{x}_n - 1, \bar{x}_n + 1] \subset \mathbb{R}^n$.
   - performs the constrained optimization with respect to the following domains:

     $$\begin{aligned} X_1 =& [\bar{x}_1 - \epsilon_1, \bar{x}_1 + \epsilon_1] \times \cdots \times [\bar{x}_{n/2} - \epsilon_1, \bar{x}_{n/2} + \epsilon_1] \times \cdots \\ & \cdots \times [\bar{x}_{n/2+1} - \epsilon_2, \bar{x}_{n/2+1} - \epsilon_2] \times \cdots \times [\bar{x}_n - \epsilon_2, \bar{x}_n - \epsilon_2] \end{aligned} \tag{1}$$

     and

     $$X_2 = B(\bar{\boldsymbol{x}} + \epsilon_1 \boldsymbol{e}\ ,\ \epsilon_2) \quad (\text{ball of center } \bar{\boldsymbol{x}} + \epsilon_1 \boldsymbol{e}, \text{ radius } \epsilon_2)\,, \tag{2}$$

     where ϵ_1 is uniformly sampled from $(0, 0.5]$ and ϵ_2 is uniformly sampled from $(0, 1]$.
   - Concerning any possible **parameter** of the methods implemented (e.g., Back-tracking parameters, stopping criteria, etc.), **tune it and discuss/motivate your choice**. If you perform preliminary tests with parameters that are not your final choice, you do not need to include the numerical results in the report, but you must briefly describe what happened in the preliminary tests.

3. [**1 point**] Extend the study in the case where derivatives are computed by using **an efficient implementation of the finite differences** (general purpose finite differences are not recommended). **Report the strategy** you adopted for implementing the finite differences in a specific section/paragraph (see Section 1.2).

   In particular, approximate the gradient considering the following values for the increment h for each differentiation:

   $$h = 10^{-k}, \qquad k = 4, 8, 12\,,$$

   and also the case of the specific increment h_i when differentiating with respect to the variable x_i, according to these values:

   $$h_i = 10^{-k}|\hat{x}_i|, \qquad k = 4, 8, 12\,, \quad i = 1, \ldots, n\,,$$

   where $\hat{\boldsymbol{x}} = (\hat{x}_1, \ldots, \hat{x}_n) \in \mathbb{R}^n$ is the point at which the derivatives have to be approximated.

4. After running the optimization methods according to the instructions above, realize a throughout comparison among the methods for each test problem (see item related to comments, Section 1.2), reporting data in tables and/or figures (see item related to tables/figures, Section 1.2). Comment on your results also in view of the expected theoretical behavior.

   For commenting the results, for each method, for each test problem, you can use **table templates** like Tables 1 and 2. You can add extra columns/information in the tables if you believe they/it can help the analysis.

   The following figures are **mandatory**:

   - top view of the function and of all the sequence paths, for the cases $n = 2$. Collect in the same figure the paths for each different starting point/simplex; create separate figures for the finite difference-based sequences. **Be sure that the figures are readable**.
   - Experimental rates of convergence for the sequences that converged. Collect in the same figure the rates different starting point; create separate figures for each dimension n, and for the finite difference-based sequences. **Be sure that the figures are readable**.

3.2 Assignment: Equality Constrained Quadratic Programming (max 10 pts)

1. [**2 points**] Consider and implement the quadratic problem

$$\begin{aligned}
\min_{x \in \mathbb{R}^n} \quad & \textstyle\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n-1} x_i x_{i+1} + \sum_{i=1}^{n} x_i \\
\text{s.t.} \quad & \text{the sum } x_1 + x_{1+K} + x_{1+2K} + \dots \text{ should be } \epsilon \\
& \text{the sum } x_2 + x_{2+K} + x_{2+2K} + \dots \text{ should be } \epsilon \\
& \vdots \\
& \text{the sum } x_K + x_{2K} + x_{3K} + \dots \text{ should be } \epsilon
\end{aligned}$$

   where $\epsilon \in (0, 1)$ is a random value generated with uniform distribution in $(0, 1)$. Generate ϵ after setting a random seed equal to the minimum student ID of the team members; e.g., given three team members with student IDs 9876, 1234, and 2345, the random seed must be set equal to $1234 = \min\{9876, 1234, 2345\}$.

2. [**8 points**] Solve the problem above following these instructions:

   - study the problem for these values: $n = 2, 2 \cdot 10^3, 2 \cdot 10^4, 2 \cdot 10^5$ with corresponding values of K that are $K = 1, 100, 500, 1000$, respectively.
   - solve the KKT conditions with the following strategies
     - full solution of the KKT system with direct solvers (whenever possible, i.e. no memory fault)
     - full solution of the KKT with a suitable iterative solver, both without and with preconditioning.
     - Schur complement approach (both with and without having the inverse of Q).
     - Null space method.

     Recommendation: Whenever you use an iterative solver for solving linear systems, report results both with preconditioning and without preconditioning.
   - For the case $n = 2$, $K = 1$, visualize the quadratic function, the constraints, and the solutions identified by all the methods.