MATH60005/70005: Optimisation (Autumn 22-23)

Instructions: read this first!

- This coursework has a total of 20 marks and accounts for 10% of the module.
- Students who want to take the final exam (90%) must submit this coursework.
- **Submission deadline:** Thursday November 24th, 13:00 UK time, via Blackboard drop box.
- Submit a single file, ideally a pdf typed in LaTeX or similar. Handwritten answers (whenever possible) are allowed but readability is essential and part of the assessment.
- **Marking criteria:** Full marks will be awarded for work that 1) is mathematically correct, 2) shows an understanding of material presented in lectures, 3) gives details of all calculations and reasoning, and 4) is presented in a logical and clear manner.
- Do not discuss your answers publicly via our forum. If you have any queries regarding your interpretation of the questions, please contact the lecturer at dkaliseb@imperial.ac.uk
- Beware of plagiarism regulations. This is an **individual assessment**.

Questions

1. **Part I: Unconstrained Optimisation**

   i) [**4 marks**] Construct a continuous function $f : \mathbb{R}^2 \to \mathbb{R}$ which is not coercive and satisfies that for any $\alpha \in \mathbb{R}$

   $$\lim_{|x_1|\to\infty} f(x_1, \alpha x_1) = \lim_{|x_2|\to\infty} f(\alpha x_2, x_2) = \infty\,.$$

   ii) [**3 marks**] Find and characterise all the stationary points of:

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

2. **Part II: Linear Least Squares**

   Consider a dynamical process of the form

   $$x_0 = \bar{x} \in \mathbb{R}\,,$$
   $$x_i = a x_{i-1} + d u_i\,, \quad i = 1, \ldots, N$$

where $a, d \in \mathbb{R}$. The variables x_i and u_i denote the internal state of the system and a control variable at discrete time i, respectively. The sequence $\mathbf{x}_{\bar{x}}^{\mathbf{u}} := \{x_i\}_{i=0}^{N} \in \mathbb{R}^{N+1}$ is the trajectory of the system departing from the initial condition $\bar{x}$ associated to the control sequence $\mathbf{u} := \{u_i\}_{i=1}^{N} \in \mathbb{R}^{N}$.

Given an initial condition $\bar{x}$, and parameters a, d, and N, our goal is to find an optimal sequence of controls $\mathbf{u}$ which drive the trajectory of the system $\mathbf{x}_{\bar{x}}^{\mathbf{u}}$ to zero while balancing the amount of control energy that is spent in this task. We express our goal as a **dynamic optimisation problem** of the form

$$\min_{\mathbf{u}\in\mathbb{R}^N} \ \|\mathbf{x}\|_2^2 + \frac{\gamma}{2}\|\mathbf{u}\|_2^2 \,, \qquad \gamma > 0 \tag{DO}$$

subject to

$$x_0 = \bar{x}\,,$$
$$x_i = a x_{i-1} + d u_i\,, \quad i = 1, \dots, N.$$

i) [**4 marks**] Reformulate (DO) as a regularised linear least squares problem for $\mathbf{u}$. Discuss existence and uniqueness of an optimal solution $\mathbf{u}^*$ to this problem. Show that any $\mathbf{u}$ solving the associated unregularised linear least squares problems (that is, with $\gamma = 0$), satisfies $\|\mathbf{u}^*\| \leq \|\mathbf{u}\|$.

ii) [**3 marks**] For $N = 40$, $a = 1$, $d = -0.01$ and $\bar{x} = 1$, generate two plots, one for the optimal control signals $\mathbf{u}^*$, and another for the associated optimal trajectories $\mathbf{x}_{\bar{x}}^{\mathbf{u}^*}$ for $\gamma = 10^{-3}, 10^{-2}, 0.1, 1$. What is the effect of increasing γ in both the control and the trajectories?

iii) [**3 marks**] In many practical applications we want to impose additional bounds on the control signal $\mathbf{u}$. For example, we want to establish an upper bound $u_i \leq u_{max}$ for all i. One way to enforce such a constraint is by using a penalty function, that is, the objective function in (DO) is replaced by

$$\min_{\mathbf{u}\in\mathbb{R}^N} \ \|\mathbf{x}_{\bar{x}}^{\mathbf{u}}\|_2^2 + \frac{\gamma}{2}\|\mathbf{u}\|_2^2 - \delta\sum_{i=1}^{N}\log(u_{max} - u_i)\,, \qquad \gamma, \delta > 0$$

Generate separate plots for the state and the control when solving the unconstrained problem (as in ii)), for $N = 40$, $a = 1$, $d = -0.01$, $\bar{x} = 1$, $u_{max} = 8$, and $\gamma = \delta = 10^{-2}$. What do you observe? You can use any method discussed in the module, but you need to state your settings.

iv) [**3 marks**] As discussed during the lectures, one may also wish to promote sparsity in the control signal $\mathbf{u}$ by considering an ℓ_1 norm penalty in the cost,

$$\min_{\mathbf{u}\in\mathbb{R}^N} \ \|\mathbf{x}_{\bar{x}}^{\mathbf{u}}\|_2^2 + \frac{\gamma_2}{2}\|\mathbf{u}\|_2^2 + \gamma_1\|\mathbf{u}\|_1 \,, \qquad \gamma_1, \gamma_2, \delta > 0\,,$$

however, we haven't discussed yet how to deal with the non-differentiability of the ℓ_1 norm at the origin. Instead, we propose the following approximations to the ℓ_1 norm:

$$\mathcal{L}_\epsilon(u_i) := \begin{cases} \frac{1}{2}u_i^2 & \text{if } \ |u_i| \leq \epsilon \\ \epsilon\left(|u_i| - \frac{1}{2}\epsilon\right) & \text{otherwise } \ . \end{cases}$$

Explain in your own words the meaning of the $\mathcal{L}_\epsilon(u_i)$ as a regulariser, is it a differentiable function? Implement a gradient descent method with backtracking -describe all your settings- to find the optimal solution to

$$\min_{\mathbf{u}\in\mathbb{R}^N} \ \|\mathbf{x}_{\bar{x}}^{\mathbf{u}}\|_2^2 + \frac{\gamma_2}{2}\|\mathbf{u}\|_2^2 + \gamma_1 \sum_{i=1}^{N} \mathcal{L}_\epsilon(u_i)\,, \qquad \gamma_1, \gamma_2, \epsilon > 0\,,$$

and compare the cases

i) $N = 40$, $a = 1$, $d = -0.01$, $\bar{x} = 1$, $\epsilon = 3$, $\gamma_2 = 10^{-2}$, $\gamma_1 = 0\,$,
ii) $N = 40$, $a = 1$, $d = -0.01$, $\bar{x} = 1$, $\epsilon = 3$, $\gamma_2 = 0$, $\gamma_1 = 10^{-2}\,$.