MATH60005/70005: Optimisation (Autumn 24-25)

Instructions: read this first![1]

- This coursework has a total of 20 marks and accounts for 10% of the module.
- **Submission deadline:** Thursday, 14th November, 13:00 UK time, via Blackboard dropbox.
- **Submit a single pdf file typeset in LaTeX or similar.**
- **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 lecturers at {dkaliseb,kloayzar}@imperial.ac.uk
- Beware of plagiarism regulations. This is a **group-based assessment** with groups from 1 to 3 students. **Make a single group submission displaying the CID of every group member in the frontpage. Do not include your name.**

Questions

Part I: Unconstrained Optimisation (5 marks)

a) Let $f(x_1, x_2) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$.

i) [**1 mark**] Is the function f coercive? Explain your answer.

ii) [**1 mark**] Find all the stationary points of f and classify them according to whether they are saddle points, strict/nonstrict local/global minimum/maximum points.

b) [**3 mark**] Let $\mathbf{d} \in \Delta_n$ with Δ_n the unit simplex. Is the matrix $\mathbf{A}$ defined by

$$A_{ij} = \begin{cases} d_i - d_i^2 & \text{if } i = j, \\ -d_i d_j & \text{otherwise,} \end{cases}$$

positive semidefinite? Justify your answer.

[1] The ability to read and follow instructions is a highly valued skill in the job market and elsewhere.

Part II: Variational Signal Denoising (10 marks)

You are given the noisy signal $\mathbf{f} \in \mathbb{R}^{1000}$ shown in Figure 1 (corresponding to the file `noisy_signal.mat`), and the goal is to use linear least squares and additional optimisation algorithms to denoise it.

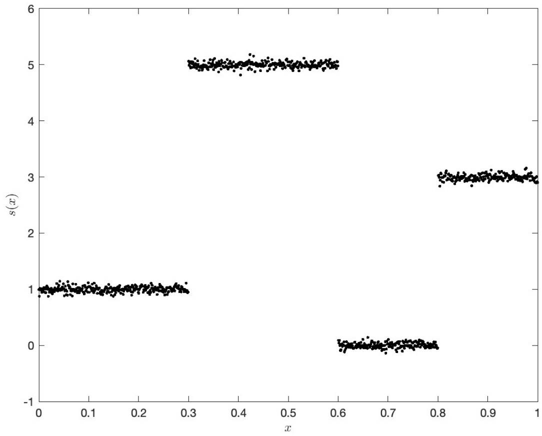

Figure 1: A piecewise constant noisy signal

One option to denoise a signal is given by the following non-smooth optimisation problem:

$$\min_{\mathbf{x}} \quad \frac{1}{2}\|\mathbf{x}-\mathbf{f}\|_2^2 + \omega\|\mathbf{L}\mathbf{x}\|_1, \qquad \omega > 0, \tag{NSD}$$

where $\|\mathbf{L}\mathbf{x}\|_1 = {}^1\!/\!_h \sum_{i=1}^{999} |x_i - x_{i+1}|$, with $h = {}^1\!/\!_{999}$. Notice that the second term is non-differentiable and therefore it restricts the kind of algorithms we can use to solve the problem.

a) **[3 marks]** As a starting point, we consider a smooth version of the problem, namely,

$$\min_{\mathbf{x}} \quad \frac{1}{2}\|\mathbf{x}-\mathbf{f}\|_2^2 + \frac{\omega}{2}\|\mathbf{L}\mathbf{x}\|_2^2, \qquad \omega > 0.$$

Write the explicit solution for this problem depending on ω and $\mathbf{f}$, evaluate it for 3 different values of $\omega \in \{10^{-4}, 5\cdot 10^{-4}, 10^{-3}\}$ and present your results.

b) To solve the original non-smooth denoising problem (NSD), we consider the following iterative algorithm:

$$(\mathbf{x}^{k+1}, \mathbf{y}^{k+1}) = \arg\min_{\mathbf{x},\mathbf{y}} \quad \frac{1}{2}\|\mathbf{x}-\mathbf{f}\|_2^2 + \frac{\lambda}{2}\|\mathbf{y}-\mathbf{L}\mathbf{x}-\mathbf{z}^k\|_2^2 + \omega\|\mathbf{y}\|_1 \tag{1}$$

$$\text{with the update} \quad \mathbf{z}^{k+1} = \mathbf{z}^k + \left(\mathbf{L}\mathbf{x}^{k+1} - \mathbf{y}^{k+1}\right). \tag{2}$$

This formulation is more convenient than (NSD) because there is no $\mathbf{L}$ inside the $\|\cdot\|_1$ norm. However, we are still dealing with a non-smooth problem.

i) **[3 marks]** We consider the following approximation of the $\|\cdot\|_1$ norm:

$$\|\mathbf{u}\|_1 \approx \|\mathbf{u}\|_{1,\gamma,\bar{\omega}} := \sum \bar{\omega}_i h_\gamma(u_i)\,,$$

where

$$h_\gamma(t) = \begin{cases} \frac{\gamma t^2}{2} & \text{if } |t| \le {}^1\!/\!_\gamma, \\ |t| - {}^1\!/\!_{2\gamma} & \text{otherwise,} \end{cases}\,, \qquad \text{for any } t \in \mathbb{R}.$$

Show that the derivative $h'_\gamma(t)$ is a continuous function and compute an expression for $\nabla\|\mathbf{u}\|_{1,\gamma,\bar{\omega}}$. With this approximation in mind, defining $\mathbf{u} = (\mathbf{x}, \mathbf{y})$, obtain an equivalent formulation of the problem (1) in the form

$$\min_{\mathbf{u}} \quad \frac{1}{2}\|\bar{\mathbf{A}}\mathbf{u} - \mathbf{f}\|_2^2 + \frac{\lambda}{2}\|\bar{\mathbf{L}}\mathbf{u} - \mathbf{z}\|_2^2 + \|\mathbf{u}\|_{1,\gamma,\bar{\omega}}\,, \tag{3}$$

giving precise definitions for $\mathbf{u}$, $\bar{\mathbf{A}}$, $\bar{\mathbf{L}}$ and $\bar{\omega}$. Solve iteratively for $\mathbf{u}$ in (3) using a gradient method with backtracking and the stopping criterion

$$\frac{\|\mathbf{u}^{k+1} - \mathbf{u}^k\|}{\|\mathbf{u}^{k+1}\|} < 10^{-4}\,,$$

or a maximum of 500 iterations. Then, update $\mathbf{z}$ as in (2), and repeat the procedure until

$$\frac{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|}{\|\mathbf{x}^{k+1}\|} < 10^{-5}.$$

Use the following values for the parameters $\lambda = 5 \cdot 10^{-4}$, $\gamma = 10^3$ and $\omega = 75$. Report the number of iterations, for both the gradient method and the full algorithm, together with the execution time.

ii) **[3 marks]** Another method is to consider the minimisation of (1) with respect to $\mathbf{x}$ and $\mathbf{y}$ separately, exploiting the fact that there exists an explicit expression for the non-smooth optimisation problem with respect to $\mathbf{y}$.
For this, let $k = 0$ and set $\mathbf{x}^k = \mathbf{f}$, $\mathbf{y}^k = \mathbf{L}\mathbf{x}^k$ and $\mathbf{z}^k = \mathbb{1}_{999}$ (vector of ones).
Repeat:

Step 1: Solve the smooth problem

$$\boxed{\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \quad \frac{1}{2}\|\mathbf{x} - \mathbf{f}\|_2^2 + \frac{\lambda}{2}\|\mathbf{y}^k - \mathbf{L}\mathbf{x} - \mathbf{z}^k\|_2^2} \tag{4}$$

Step 2: Update the solution of the non-smooth problem

$$\mathbf{y}^{k+1} = \arg\min_{\mathbf{y}} \quad \frac{\lambda}{2}\|\mathbf{y} - \mathbf{L}\mathbf{x}^{k+1} - \mathbf{z}^k\|_2^2 + \omega\|\mathbf{y}\|_1,$$

given by:

$$y_i^{k+1} = \mathcal{S}\left(\mathbf{L}_i\mathbf{x}^{k+1} + z_i^k, {}^\omega\!/\!_\lambda\right) \qquad \forall i = 1, \ldots, 999$$

where $\mathbf{L}_i$ denotes the i-th row of the matrix $\mathbf{L}$ and $\mathcal{S}$ acts component-wise and is given by the following expression:

$$\mathcal{S}\,(t, \eta) = \frac{t}{|t|}\max\{|t| - \eta, 0\} \qquad \text{for any } t \in \mathbb{R} \text{ and any } \eta > 0.$$

Step 3: Update $\mathbf{z}^{k+1} = \mathbf{z}^k + (\mathbf{L}\mathbf{x}^{k+1} - \mathbf{y}^{k+1})$, and return to *Step 1* unless a stopping criterion is satisfied.

Find the explicit solution to problem (4) and solve the denoising problem iterating over *Steps 1- 3* using the parameters $\lambda = 10^{-2}$, $\omega = 2 \cdot 10^{-3}$, and stopping criterion:

$$\frac{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|}{\|\mathbf{x}^{k+1}\|} < 10^{-5}.$$

Report the number of iterations and time of execution for this algorithm.

c) [**1 mark**] Discuss the differences you observe in the results comparing the three proposed methods in a), bi), and bii).

Part III: Choose only one of the following two questions. If you submit more than one answer, only the first answer will be marked.

Option A) **Gradient Descent (5 marks)**

Let $g \colon \mathbb{R}^3 \to \mathbb{R}$ given by

$$g(\mathbf{x}) = \frac{1}{4}\sum_{i=1}^{2}\cos(x_i - x_{i+1}) + \sum_{i=1}^{3} i x_i^2$$

i) [**4 marks**] Compute a fixed step-length for which the gradient method is guaranteed to converge. *Hint: use that for a matrix* $\mathbf{A}$, $\|\mathbf{A}\| \leq \|\mathbf{A}\|_F$.

ii) [**1 marks**] Find one stationary point of g and classify it according to whether it is saddle point, strict/nonstrict local/global minimum/maximum point. Justify your answer.

Option B) **Variational Image Denoising (5 marks)**

You are given the noisy image $\mathbf{f} \in \mathbb{R}^{100\times 100}$ shown in Figure 2 (corresponding to the file `noisy_image.mat`). The goal is to use the previously proposed algorithms to denoise it.

As in the signal denoising case, we propose the following nonlinear optimisation problem:

$$\min_{\mathbf{x}} \quad \frac{1}{2}\|\mathbf{x} - \mathbf{f}\|_2^2 + \omega\|\mathbf{D}\mathbf{x}\|_1 \qquad \omega > 0,$$

where $\mathbf{D}$ represents a dicretised 2-dimensional difference operator. The second term can be rewritten as $\|\mathbf{D}\mathbf{x}\|_1 = \|\mathbf{D}_1\mathbf{x}\|_1 + \|\mathbf{D}_2\mathbf{x}\|_1$, where $\mathbf{D}_1, \mathbf{D}_2$ represent the horizontal and vertical discrete difference operators, respectively. The MATLAB code to generate these operators is described in what follows:

```
[nx1,nx2] = size(f_noisy);
nx = nx1 * nx2;
h1=1/(nx1-1);
h2=1/(nx2-1);
e = ones(nx,1);
% Horizontal difference operator
```

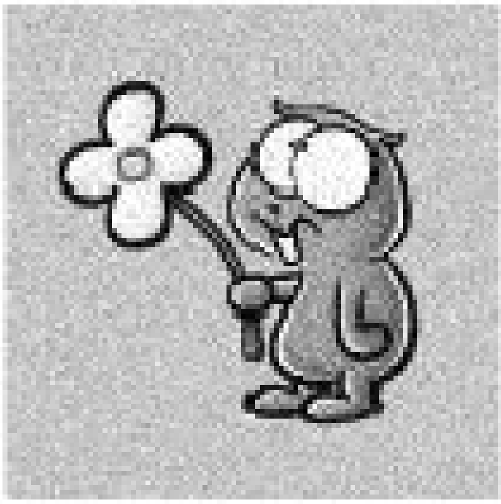

Figure 2: A black&white noisy image

```
D1 = (1/h1)*spdiags([-e,e],[0,nx1],nx1*(nx2-1),
   nx);
% Vertical differential operator
D2row = spdiags([-e,e],[0,1],nx1-1,nx1);
D2 = (1/h2)*kron(speye(nx2,nx2),D2row);
D=[D1;D2];
```

Images are typically represented as matrices, where each pixel corresponds to a matrix entry containing its color intensity. Since we need vectors to define our optimisation problem, we use the vectorisation of a matrix. This can be represented as $\mathbf{f} = \mathrm{vec}(\mathbf{f}) \in \mathbb{R}^{10000}$. In MATLAB, vectorisation can be achieved with the command `f_noisy(:)`. Once we obtain a solution, we need to reshape it back into a matrix form to display it. In MATLAB, this can be done with the command `imshow(uint8(reshape(solution, nx1, nx2)))`, where `nx1` and `nx2` are the original dimensions of the image.

i) [**2 marks**] Consider the problem

$$\min_{\mathbf{x}} \quad \frac{1}{2}\|\mathbf{x} - \mathbf{f}\|_2^2 + \frac{\omega}{2}\|\mathbf{D}\mathbf{x}\|_2^2$$

Find an explicit solution for this problem and report the results obtained with $\omega = 10^{-4}$.

ii) [**3 marks**] Solve the problem using the algorithm described in Part II, question b ii). Use the parameters $\lambda = 10^{-4}$, $\omega = 0.09$, and iterate over *Steps 1-3* until the stopping criterion:

$$\frac{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|}{\|\mathbf{x}^{k+1}\|} < 10^{-4},$$

is satisfied. Report the results and comment on the difference between the obtained solution and the one from part i).