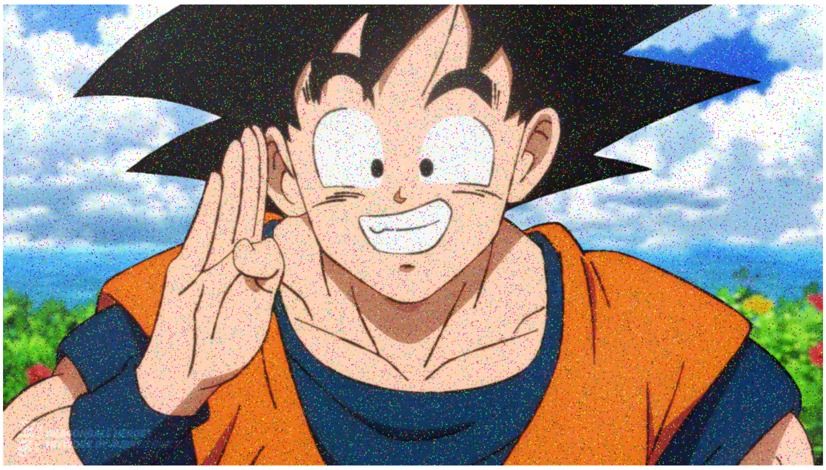
LINMA 2471 – Optimization Models and Methods II

Homework II

Projected Gradient Method

1. Consider the following noisy image of Son Goku (file **son_goku.png**):



Let $Y^{(1)}$,$Y^{(2)}$,$Y^{(3)}$$\in [0,255]^{m\times n}$ be the three pixel matrices that specify this noisy image in the primary colors. One approach to take out the noise from the image consists in solving the optimization problems:

$$\begin{aligned} \min_X \quad & f_\ell(X) \equiv \tfrac{1}{2}\|X - Y^{(\ell)}\|_F^2 + \lambda R(X), \\ \text{s.t.} \quad & X \in [0,255]^{m\times n}, \end{aligned} \tag{1}$$

for $\ell = 1,2,3$, where $\lambda > 0$ and

$$R(X) = \sum_{i=2}^{m-1}\sum_{j=2}^{n-1} (X_{i+1,j} - X_{i,j})^2 + (X_{i,j+1} - X_{i,j})^2 \, .$$

The solutions $X_*^{(1)}$, $X_*^{(2)}$ and $X_*^{(3)}$ specify the *denoised image.*

(a) Compute $\nabla f_\ell(\cdot)$. Provide the detailed derivation.

(b) Show that $\nabla f_\ell(\cdot)$ is L-Lipschitz and give the constant L explicitly.

(c) Is $f_\ell(\cdot)$ convex? Prove your answer.

(d) Implement the Projected Gradient Method to solve (1) with stopping criterion $\|G_L(X_k)\|_F \leq \epsilon$.

(e) Apply your code with $\epsilon = 10^{-5}$ to compute approximations to $X_*^{(1)}$, $X_*^{(2)}$ and $X_*^{(3)}$ considering four different choices for λ. For each λ, form the denoised image and provide the corresponding file in format **.png**.

(f) What is the relation between the *quality* of the denoised image and the magnitute of the regularization parameter λ?

Proximal Gradient Method

2. Among different types of mathematical models, differential equations constitute a powerful tool for modelling the evolution of systems over time. The derivation of the differential equations for an specific phenomenon is usually done from first principles. Recent advances in Machine Learning have motivated several attempts to automate the discovery of differential equations directly from data available about the phenomenon. In this direction, a promising data-driven approach is the Sparse Identification of Nonlinear Dynamics (SINDy). For a given phenomenon, the SINDy approach seeks a system of Ordinary Differential Equations (ODEs) of the form

$$\begin{array}{rcl} \dot{x}_1(t) & = & f_1(x_1(t),\ldots,x_n(t)), \\ \vdots & \vdots & \vdots \\ \dot{x}_n(t) & = & f_n(x_1(t),\ldots,x_n(t)), \end{array} \tag{2}$$

where $x_\ell : \mathbb{R}_+ \to \mathbb{R}$ represents the ℓ-th state at time t, and $f_\ell : \mathbb{R}^n \to \mathbb{R}$ $(i = 1,\ldots,n)$ are the unknown functions that define the laws that govern the phenomenon. In order to identify $f_\ell(\,\cdot\,)$, the first step is to assume that

$$f_\ell(x_1,\ldots,x_n) = \sum_{j=1}^{p} \alpha_j \theta_j(x_1,\ldots,x_n), \tag{3}$$

where $\theta_j : \mathbb{R}^n \to \mathbb{R}$ $(i = 1,\ldots,p)$ are candidate functions. The second step is to obtain samples of the ℓ-th state at times $t_1,\ldots,t_m$ and, from this data, approximate the derivatives $\dot{x}_\ell(t_i)$ by forward/backward finite differences:

$$b_i^{(\ell)} = \begin{cases} \dfrac{x_\ell(t_{i+1}) - x_\ell(t_i)}{t_{i+1} - t_i}, & i = 1,\ldots,m-1, \\[2ex] \dfrac{x_\ell(t_i) - x_\ell(t_{i-1})}{t_i - t_{i-1}}, & i = m, \end{cases} \tag{4}$$

Then, in view of (2), (3) and (4), the identification of $f_\ell(\,\cdot\,)$ reduces to the problem of finding $\alpha \in \mathbb{R}^p$ such that

$$b_i^{(\ell)} \approx \sum_{j=1}^{p} \alpha_j \theta_j(x_1(t_i),\ldots,x_n(t_i)), \quad i = 1,\ldots,m.$$

To favor the discovery of a parsimonius model, SINDy computes α by solving the optimization problem

$$\min_{\alpha \in \mathbb{R}^p} \frac{1}{2}\|b^{(\ell)} - \Theta\alpha\|_2^2 + \lambda\|\alpha\|_1, \tag{5}$$

where $\lambda > 0$, $b^{(\ell)} = \begin{bmatrix} b_1^{(\ell)} & \dots & b_m^{(\ell)} \end{bmatrix}^T$, and $\Theta \in \mathbb{R}^{m \times p}$ with

$$\Theta_{ij} = \theta_j(x_1(t_i), \dots, x_n(t_i)).$$

Denoting by $\alpha^{(\ell)}$ the compute (approximate) solution of (5), the identified system of ODEs is given by

$$\begin{array}{rcl} \dot{x}_1(t) & = & \sum_{j=1}^p \alpha_j^{(1)} \theta_j(x_1(t), \dots, x_n(t)), \\ \vdots & \vdots & \vdots \\ \dot{x}_n(t) & = & \sum_{j=1}^p \alpha_j^{(n)} \theta_j(x_1(t), \dots, x_n(t)). \end{array} \tag{6}$$

In this exercise you will test if SINDy is able to identify the SIR model for a pandemic scenario:

$$\begin{array}{rcl} \dot{x}_1(t) & = & -\beta x_1(t) x_2(t), \\ \\ \dot{x}_2(t) & = & \beta x_1(t) x_2(t) - \gamma x_2(t), \\ \\ \dot{x}_3(t) & = & \gamma x_2(t), \end{array} \tag{7}$$

where $x_1(t)$, $x_2(t)$ and $x_3(t)$ represent, respectively, the percentage of the population that is susceptible, infected and recovered at time t.

(a) Considering $\beta = 0.2$, $\gamma = 0.05$, $x_1(0) = 0.995$, $x_2(0) = 0.005$ and $x_3(0) = 0$, solve numerically system (7) for $t_i = i - 1$ and $i = 1, \dots, 201$. Form the matrix $X \in \mathbb{R}^{201 \times 3}$ with

$$X_{i\ell} = x_\ell(t_i), \quad i = 1, \dots, 201, \quad \ell = 1, 2, 3,$$

and compute the corresponding vectors $b^{(1)}$, $b^{(2)}$ and $b^{(3)}$ by (4). Provide the code used to solve (7). Also, in a **.txt** file, provide the data matrix X and also the vectors $b^{(1)}$, $b^{(2)}$ and $b^{(3)}$.

(b) Implement the Proximal Gradient Method to solve (5) for generic matrix $\Theta \in \mathbb{R}^{m \times p}$ and vector $b \in \mathbb{R}^m$ with stopping criterion $\|G_L(\alpha)\|_2 \leq \epsilon$, where

$$G_L(\alpha) = L \left(\alpha - \text{prox}_{\frac{\lambda}{L}\|\cdot\|_1} \left(\alpha - \frac{1}{L} \nabla f(\alpha) \right) \right),$$

with $f(\alpha) = \frac{1}{2}\|b - \Theta\alpha\|_2^2$ and L beign a Lipschitz constant of $\nabla f(\,\cdot\,)$. Provide the code of your implementation.

(c) Considering the candidate functions

$$\theta_j(x_1, x_2, x_3) = \begin{cases} x_1, & \text{if } j = 1, \\ x_2, & \text{if } j = 2, \\ x_3, & \text{if } j = 3, \\ x_1 x_2, & \text{if } j = 4, \\ x_2 x_3, & \text{if } j = 5, \end{cases}$$

apply SINDy to identify the system of ODEs from the data collected at item (a), using $\epsilon = 10^{-5}$ and $\lambda = 0.001$. Report the obtained vectors $\alpha^{(1)}$, $\alpha^{(2)}$ and $\alpha^{(3)}$, and the identified system (6). Does SINDy works?

(d) Try at least three different values of $\lambda > 0$. How changes in λ affect the output of SINDy?