Signal Processing and Optimization for Big Data

17/10/2025

1) Let us consider the LASSO problem

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \alpha\|\mathbf{x}\|_1, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{N \times M}$, $\mathbf{x} \in \mathbb{R}^{M \times 1}$ and $\alpha > 0$. Write two functions, using a programming language of your choice (e.g., Matlab/Python) to solve the problem through **ISTA** and **ADMM**.

2) Test the functions considering a Gaussian random matrix $\mathbf{A}$ and generating $\mathbf{y}$ from a 3-sparse vector $\mathbf{x_0}$. Let K be the number of **iterations**

- Plot the temporal behavior of the reconstruction MSE $E[k] = \|\mathbf{x}[k] - \mathbf{x_0}\|^2$ and the objective function $J[k]$.

- Plot the sparsity of the obtained solution and $\mathbf{x}[K]$ and the final reconstruction error $E[K]$ for different values of α.

- Plot the temporal behavior of the two terms of the cost function for different values of α.

Hint: Use, for example, N=10, M=20, K=200.

3) Let us consider the following linear observation model

$$\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \mathbf{n}, \tag{2}$$

where $\mathbf{y} \in \mathbb{R}^N$, $\mathbf{A} \in \mathbb{R}^{N \times M}$ and $\mathbf{n} \sim \mathcal{N}(0, \sigma_n^2 \mathbf{I})$. Let us also consider the Ridge regression problem

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \alpha\|\mathbf{x}\|_2, \tag{3}$$

- Write a function to solve the Ridge regression problem trough an iterative algorithm.

- Fix a noise variance σ_n of your choice, and generate some synthetic data according to the linear model (2). Choose $\mathbf{A}$ as a random matrix, and assume $\mathbf{x}_0 \sim \mathcal{N}(0, \sigma_0 \mathbf{I})$.

- Plot the temporal behaviour of the cost function $J[k]$ for different values of the regularization parameter α and compare the convergence speed of Ridge regression solved through GD and LASSO solved through the ISTA method.

- Considering the same observation model, solve the LASSO problem for different values of the regularization parameter α and plot: i) the temporal behaviour of the cost function and the reconstruction MSE for different values of α; ii) the sparsity of the solution as a function of the parameter α; .

4) Let us consider the *mpg-dataset*. This data-set contains **398** instances of car models for which we want to predict the fuel consumption $\mathbf{y}$, expressed in miles per gallon (MPG). We have 7 features, used to build the design matrix $\mathbf{A}$. Choose one method between ISTA and ADMM and apply it to solve this regression problem.

- Select the best trade-off parameter α through cross-validation, considering N_{tr} samples for training and N_{val} samples for validation.

- For each value of α, evaluate the sparsity of the obtained solution and the generalization capabilities of the model.

- Plot the temporal behavior of the terms of the cost function for different values of α.

Hint 1: A possible way to evaluate a regression model is given by the R2 score, which is defined by

$$R2 = 1 - \frac{RSS}{TSS}, \tag{4}$$

where $RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ is the Residual Sum of Squares, while $TSS = \sum_{i=1}^{n}(y_i - \overline{y})^2$ is the Total Sum of Squares. Specifically, $\mathbf{y}, \hat{\mathbf{y}} \in \mathbb{R}^{N_{\text{val}}}$ are the observations and the predictions vectors respectively, while $\overline{y}$ is the mean of the observations.
Hint 2: it could be useful to normalize the data to help the convergence of the algorithms.

5) Consider a network composed of N receivers and a target t located at an unknown position $\mathbf{x}_0 \in \mathbb{R}^2$. Receiver i is at a known position $\mathbf{p}_i = (x_i, y_i)^\top$, and measures a noisy estimate of its distance from the target:

$$d_i = \left\|\mathbf{x}_0 - \mathbf{p}_i\right\| + v_i, \qquad i = 1, \ldots, N, \tag{5}$$

where v_i is additive noise (assumed zero-mean).
The goal is to estimate the target position by solving the following problem

$$\min_{\mathbf{x} \in \mathbb{R}^2} \frac{1}{2} \sum_{i=1}^{N} \left(d_i - \left\|\mathbf{x} - \mathbf{p}_i\right\| \right)^2. \tag{6}$$
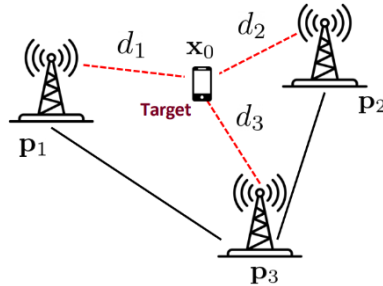

Figure 1: Considered network topology.

- Design a distributed consensus algorithm to solve the optimization problem in (6).

- Plot the temporal behaviour of (i) the disagreement among nodes and (ii) the MSE between the current estimate and the true target position $\mathbf{x}_0$.

2