Linear Regression

Lesson 2 : Lab Session

Advanced Machine Learning, CentraleSupelec

Teacher's Assistant: Omar CHEHAB

Professors : Emilie CHOUZENOUX, Frederic PASCAL

CentraleSupélec | université PARIS-SACLAY

General Information

- **Assignment :** alone or in pairs, you will code the algorithms you learnt in 'scikit-learn formalism', and apply them to images and text.

- **Due :** the 5 lab assignments for lessons 3-7 are due a week from when they are given, at aml.centralesupelec.2020@gmail.com *(today's lab is not due)*

- **Grading** : each assignment is worth 4 points — your 4 best labs out of the 5 will be retained and will count for half of your final grade.

- **Questions :** questions or feedback are welcome after class or by email at l-emir-omar.chehab@inria.fr

Lesson: recap

	Objective	**Solution:** closed-form	**Solution:** via algorithm	**Effect**
LS	$\min_\beta \|\|y - X\beta\|\|_2^2$	$\hat{\beta} = (X^T X)^{-1} X^T y$	gradient descent	vanilla
Ridge	$\min_\beta \|\|y - X\beta\|\|_2^2 + \lambda \|\|\beta\|\|_2^2$	$\hat{\beta} = (X^T X + \lambda I)^{-1} X^T y$	gradient descent	shrinkage
LASSO	$\min_\beta \|\|y - X\beta\|\|_2^2 + \lambda \|\|\beta\|\|_1^1$	unavailable	proximal gradient (ISTA): $\beta_{t+1} = ST_{\lambda\eta} \circ \left(\beta_t + \eta \cdot X^T(y - X\beta)\right)$	sparsity
Robust	$\min_\beta \rho(y - X\beta)$ e.g. $\rho = \|\|.\|\|_1^1$	unavailable	IRLS: (cf. cours) $D_t = Diag(\omega(y - X\beta_t))$ $\beta_{t+1} = (X^T D_t X)^{-1} X^T D_t y$	robustness

Assignment: plan

1. Simple Linear Regression

2. Adding complexity via a **nonlinear** ‘basis’

3. Managing complexity via **regularization** (L1, L2)

4. Incorporating robustness via the **loss function**

5. Application : predicting Bicycle Traffic