Inside Rijndael

Understanding of the arithmetic in Rijndael's finite field

Fernando Crema Garcia

Data Privacy and Security
Laurea Magistrale in Data Science
Sapienza, University of Rome

A. Y. 2019 - 2020

Table of contents

Outline

Code snippet

```c
typedef unsigned char byte;

byte mult(byte in_1, byte in_2){
  byte mask=0x01, result=0x00, piv=in_1;

  for(int i=0; i<8; i++){
    if(in_2 & mask) result^=piv;
    mask = mask << 1;
    piv = xtime(piv);
  }

  return result;
}
```

Listing 1: Mult example

Assumptions

Let us consider the usual linear regression model defined as:

$$y = \beta_*^t \mathbf{x} + \epsilon \text{ with}$$

i. $\mathbf{x} \in \mathbb{R}^p$ a vector of random variables normally called the input vector.
ii. $\epsilon \in \mathbb{R}$ the random noise defined as a Gaussian random variable with expectation zero and variance σ^2, this is, $\epsilon \sim \mathcal{N}(\mu = 0, \sigma^2)$
iii. $y \in \mathbb{R}$ a random variable that depends linearly on $\mathbf{x}$.
iv. $\beta_* \in \mathbb{R}^p$ is the optimal model.

Problem Formulation

The problems we need to solve to estimate β_* are written as follow:

$$
\begin{aligned}
P(s,\lambda) \quad & \min_{\beta, e} && f(e) + \lambda g(\beta) && && \text{(1a)}\\
& \text{s.t.} && e = \mathbf{y} - X\beta && e \in \mathbb{R}^m, \beta \in \mathbb{R}^p, && \text{(1b)}\\
& && h(\beta) \le s && s \in \mathbb{R}, s \ge 0, && \text{(1c)}\\
& && \mathbf{L} \le A\beta \le \mathbf{U} && \mathbf{L}, \mathbf{U} \in \mathbb{R}^m, A \in \mathbb{R}^{qxp} && \text{(1d)}
\end{aligned}
$$

Examples

- f is the error function. Examples: $f(\mathbf{e}) = \|\mathbf{e}\|_1$ y $f(\mathbf{e}) = \frac{1}{2}\|\mathbf{e}\|_2^2$, among others.
- h and g are the complexity functions of the model. Examples: $g(\beta) = \|\beta\|_1$ y $g(\beta) = \|\beta\|_2^2$, $h(\beta) = \|\beta\|_1$ y $h(\beta) = \|\beta\|_0 = |\{j : \beta_j \neq 0, \; j \in [n]\}|$, among others.
- A,$\mathbf{L}$ and $\mathbf{U}$ allow the modelling of linear constraints over the regressors.

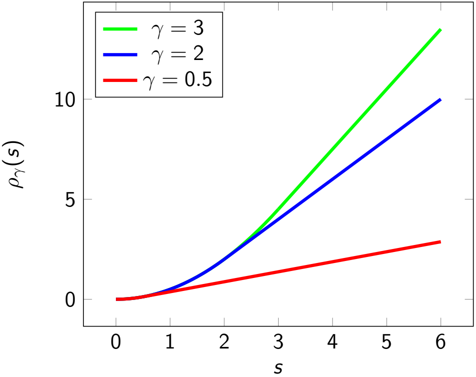
The Holistic Regression problem

$$
\begin{gathered}
(R(\lambda, s)) \ \min_{\beta,\mathbf{e},\mathbf{z}} \ f(\mathbf{e}) + \lambda g(\beta) \ \ s.t. \\
\mathbf{y} - X\beta = \mathbf{e} \\
h(\beta) \leq s \\
\mathbf{L} \leq A\beta \leq \mathbf{U} \\
(\beta, \mathbf{z}) \in H \\
\beta \in \mathbb{R}^n, \ \ \mathbf{e} \in \mathbb{R}^m, \ \ \mathbf{z} \in \{0,1\}^n
\end{gathered}
$$

- The structure of $R(\lambda, s)$ is too general, so the algorithms designed for $P(s, \lambda)$ cannot be used (Specially because H).
- For the usual optins of f,g,h and for H defined as affine equations and inequations in $(\beta, \mathbf{z})$ $R(\lambda, s)$ is as (0-1-MICQP).

Huber Function $\rho_\gamma(s)$

$$\rho_\gamma(s) = \begin{cases} \frac{1}{2}s^2 & 0 \leq s \leq \gamma \\ \gamma s - \frac{1}{2}\gamma^2 & s \geq \gamma \end{cases}$$



The ϵ-insensitive Huber Function $g_\gamma^\epsilon(t)$

$$f_\gamma^\epsilon(\mathbf{e}) = \sum_{i\in[m]} g_\gamma^\epsilon(\bar{\mathbf{e}}_i) \text{ with:}$$

$$g_\gamma^\epsilon(t) = \begin{cases} 0 & si \quad |t| \leq \epsilon \\ \rho_\gamma(|t| - \epsilon) & si \quad |t| \geq \epsilon \end{cases}$$

Outline

1. Introduction
   - Original problem
   - The Holistic Regression Problem
2. The study case
3. Roadmap

Study case

$$(R(\lambda, k)) \ \ \underset{\beta,\mathbf{e},\mathbf{z}}{min} \ \ f^{\epsilon}_{\gamma}(\mathbf{e}) + \lambda \| \ \beta \ \|_1 \ \ s.t. \ \ (1)$$

$$\mathbf{y} - X\beta = \mathbf{e} \ \ (2)$$

$$\text{If} \ \ \mathbf{z}_j = 1 \ \ \text{then} \ \ \beta_j = 0 \ \ \forall j \in [n] \ \ (3)$$

$$Co(\mathbf{z}) \leq k \ \ (4)$$

$$\sum_{j \in J1_i} \mathbf{z}_j \leq 1 \ \ \forall i \in [n_1] \ \ (5)$$

$$\sum_{j \in J2_i} \mathbf{z}_j = 1 \ \ \forall i \in [n_2] \ \ (6)$$

$$\mathbf{z}_{j_1} = \mathbf{z}_{j_2} \ \ \forall j_1, j_2 \in B_i \ \ \forall i \in [n_B] \ \ (7)$$

$$\beta_j \geq 0 \ \ \forall j \in J^+, \ \ \beta_j \leq 0 \ \ \forall j \in J^- \ \ (8)$$

Study case (II)

$$(R(\lambda, k)) \ \min_{\beta, \mathbf{e}, \mathbf{z}} \ f_{\gamma}^{\epsilon}(\mathbf{e}) + \lambda \| \beta \|_1 \ \ s.t. \ \ (1)$$

$$\ldots$$

$$\mathbf{z}_{j_1} + \mathbf{z}_{j_2} \leq 1 \ \ \forall (j_1, j_2) \in Jc \ \ (9)$$

$$\beta \in \mathbb{R}^n, \ \ \mathbf{e} \in \mathbb{R}^m, \ \ \mathbf{z} \in \{0, 1\}^n \ \ (10)$$

Outline

Road

Objective: Find a set of solutions with values close to the optimal, this is *a path of Approximate Solutions*

1. Formulation of $R(\lambda, k)$.
2. Obtaining valid values for BigM values. BigMs too big $\implies$ numerical problems, inefficiency of algorithms. BigMs too small *implies* good solutions removed.
3. *Local Holistic Searches*: Distances, Neighborhoods and Complexity function are presented on $\mathbf{z}$, which we will call Holistic, that adapt naturally to the case study, to obtain solutions locally optimal.
4. Algorithms for the construction of the Approximate Solutions Path combining (2.) and (3.)

```c
typedef unsigned char byte;

byte mult(byte in_1, byte in_2){
  byte mask=0x01, result=0x00, piv=in_1;

  for(int i=0; i<8; i++){
    if(in_2 & mask) result^=piv;
    mask = mask << 1;
    piv = xtime(piv);
  }

  return result;
}
```

Listing 2: Mult example