CS771 Assignment 1

Yogesh Dudi (201164) Aryan Srivastava (210204)
Anshul Singh Baghel (200155) Priyanshu Nain (200736)
Tarun Beniwal (201046)

Abstract

Sparse Challenge-Response Pair (CRP) Generation Physical Unclonable Function (CDU PUF) can be broken by a single sparse linear model. More specifically, give derivations for a map $\phi : \{0,1\}^D \to \mathbb{R}^D$ mapping D-bit $0/1$-valued challenge vectors to D-dimensional feature vectors and show that for any S-sparse CDU PUF, there exists an S-sparse linear model, i.e., $w \in \mathbb{R}^D$ such that $\|w\|_0 \leq S$, i.e., w has at most S non-zero coordinates, and that for all challenges $c \in \{0,1\}^D$, the following expression $w^\top \phi(c)$ gives the correct response. Note that no bias term (not even a hidden one) is allowed in the linear model.

1 Solution of Task 1

We have CRP (c,r) where c is D bit 0 or 1valued vector and r is corresponding generated response by S-sparse CRP PUF.

To Show: There exits s sparse linear model, $w \in \mathbb{R}^D$, which has most S non-zero coordinates, i.e., $\|w\|_0 \leq S$, such that for any challenge vector c, $w^\top \phi(c)$ provide correct response.

For map $\phi : \{0,1\}^D \to \mathbb{R}^D$, transforms D - bit challenge vector D dimentional feature vector.

Here mapping ϕ is a follow,
$\phi(c) = (\phi_0(c), \phi_1(c), \phi_2(c).....\phi_D(c))$

For each feature $\phi_i(c)$in teasure vector, if ith bit of c is 0 then $\phi_i(c)$ is -1, fr c equal to 1,$\phi_i(c)$ is 1.

Define, for e.g. c=(1, 0, 1, 1)
$\phi_1(c) = (-1)\hat{}(c_1) = 1$
$\phi_2(c) = (-1)\hat{}(c_2) = -1$
$\phi_3(c) = (-1)\hat{}(c_3) = 1$
$\phi_4(c) = (-1)\hat{}(c_4) = 1$
feature vector $\phi(c)$ would be (1, -1, 1,1) for c=(1, 0, 1, 1)

Since, w is S-sparse, i.e., it has most S non-zero coordinates. Let assume we have sorted list of indices $(i_1, i_2, i_3, ...i_s)$ representing S non zero coordinates of w. For to provide correct reponse, we need values of w at these indices.
we have,
$\phi_i(c) = (-1)\hat{}(c_i)$ we want,
Corresponding coordinates of w to be either -1 or +1 depending on the response of PUF for every particular challenge c.
We have two response,
if response r=0, set $w_{i,k}$ = -1 for all k from 1 to s

Preprint. Under review.

if response r=1, set $w_{1,k}$ = +1 for all k from 1 to s
When r=0,
$w^\top \phi(c)$ = (-1)$w_{i,1}\phi_{i,k}(c)$ + (-1)$w_{i,2}\phi_{i,2}(c)$ (-1)$w_{i,3}\phi_{i,3}(c)$ ..... (-1)$w_{i,s}\phi_{i,s}(c)$
$w^\top \phi(c)$ =$(-1)(-1)\hat{}c_{i,1}$ + $(-1)(-1)\hat{}c_{i,2}$ + $(-1)(-1)\hat{}c_{i,3}$ .....$(-1)(-1)\hat{}c_{i,s}$
since, PUF response is 0.
sum of $(-1)\hat{}c_{i,k}$ will be 0 for k 1 to S.
Therefore, $w^\top \phi(c)$ = 0, which matches the correct response.

similarly, when r=1,
$w^\top \phi(c)$ =$(-1)\hat{}c_{i,1}$ + $(-1)\hat{}c_{i,2}$ + $(-1)\hat{}c_{i,3}$ .....$(-1)\hat{}c_{i,s}$
since, PUF response is 1.
sum of $(-1)\hat{}c_{i,k}$ will be 1 for k 1 to S.
Therefore, $w^\top \phi(c)$ = 1, which matches the correct response.

By making the mapping ϕ and choosing the appropriate non zero coordinates for sparse linear model w, we showed that for any S sparse CDU PUF, there exists ab sparse liner model w, such that $\|w\|_0 \leq S$ and $w^\top \phi(c)$ gives correct response for any challenge c.

2 Solution of Task 2

We have submitted python

3 Solution of Task 3

The submitted my_fit() method does an iterative process of optimising the weights with hard thresholding. It tries to find a sparse answer by making some weights exactly zero and estimating others using least squares regression. The process keeps going until either the maximum number of rounds is reached or the change in weights is less than a certain tolerance.

Before we came up with the final answer, we tried the following:

Lasso regression with soft thresholding method, the following is the resulting metrics,

MSE: 1515.6251601503686
r2 score: 0.8149623043059591
MAE: 31.33836896741733

When we looked at the results, we chose to use our method, which gives,

MAE: 3.0453068288119978
MSE: 14.148352275348516
r2: 0.9989072632131443

Better than the types we had tried before.

4 Solution of Task 4

In this code, there are few hyperparameters that potentially affect performance of myfit().
They are:

1. **Learning rate:** Determines step size of each iteration of gradient descent. Too large can cause algorithm to overshoot the optimal solution, too small can result in slow convergence.

2. **Max_iteration:** Determines how long algorithm will run before terminating. Higher the Max_iteration, better the convergence, but can also increase computation time.

3. **tol:** Determines tolerance or threshold for convergence, it is the mininum change in weights required to consider the algorithm as a converged. .

To find optimal solution we used grid_search using cross validation =5 and negative mean squared error as scoring metric. Grid Search looks for combination of hyperparameter which yields lowest mean squared error.
parameter_grid used:
{
'Learning rate': [0.1, 0.2, 0.3],
'Max_iteration': [20,22,25],
'tol':[$1e^{-3}$,$1e^{-4}$,$1e^{-5}$]
}