THE NATIONAL UNIVERSITY
of SINGAPORE

School of Computing
Computing 1, 13 Computing Drive, Singapore 117417

TRB6/18

Differential Privacy for Regularised Linear Regression

Ashish Dandekar, Debabrota Basu and Stéphane Bressan

June 2018

T e c h n i c a l R e p o r t

Foreword

This technical report contains a research paper, development or tutorial article, which has been submitted for publication in a journal or for consideration by the commissioning organization. The report represents the ideas of its author, and should not be taken as the official views of the School or the University. Any discussion of the content of the report should be sent to the author, at the address shown on the cover.

Mohan KANKANHALLI
Dean of School

Differential Privacy for Regularised Linear Regression

Ashish Dandekar, Debabrota Basu, and Stéphane Bressan

School of Computing, National University of Singapore, Singapore.
(ashishdandekar, debabrota.basu)@u.nus.edu, steph@nus.edu.sg

Abstract. Recent attacks on machine learning models such as membership inference attacks increase the concern for privacy. Linear regression is such an essential statistical machine learning model at risk. For a given dataset, linear regression determines the parameters of the linear equation connecting the predictor variables to the response variable. As such linear regression yields a set of unstable and overfitted parameters. Regularisation terms are added to the loss function of linear regression in order to avoid overfitting. LASSO, ridge, and elastic net are three variants of regularised linear regression. We present an ϵ-differentially private functional mechanism for the aforementioned variants of regularised linear regression. We empirically and comparatively analyze its effectiveness. A functional mechanism achieves differential privacy for linear regression by adding noise to the loss function. We empirically show that an ϵ-differentially private functional mechanism causes more error than the non-private linear regression models whereas their performances are comparable. We also discuss caveats in the functional mechanism, such as non-convexity of the noisy loss function, which causes instability in the results of differentially private linear regression models. This discussion puts forth the need of designing a differentially private mechanism that produces a noisy loss function that is convex.

Keywords: linear regression, data privacy, differential privacy, elastic net

1 Introduction

Recent attacks on machine learning models increase the concern for the privacy of users and promotes a need to protect it [9]. Different attack models, such as the membership inference attack [19] and the white-box and the black-box attacks [20] are designed and studied by researchers to explore and to expose the vulnerability of machine learning models. Leveraging the knowledge of such attacks, a malevolent data analyst can breach the privacy of a machine learning model by inferring the values of some attributes of the data points that are used to train the model [24]. Specifically under the Machine Learning as a Service (MLaaS) model [18] the user cannot control the data leak caused by the machine learning model catered by the service provider [20]. Therefore, there is a need

for privacy guarantees for machine learning models, especially when they are trained on the data that contains highly sensitive attributes.

Dwork et al. proposed a probabilistic framework, called *differential privacy* [8], to quantify privacy (Section 3.1). Differential privacy provides a framework to design privacy inducing mechanisms [7, 26] that introduce noise at different stages of a machine learning model, for instance in the output of the model [7] or in the input of the model [15], to make it as differentially private as required.

In this paper, we study differential privacy inducing mechanisms for *linear regression* models under the release of the model itself. Linear regression models constitute a family of widely used machine learning models in such eclectic domains as econometrics [14], medicine [2, 21] and policy making [4]. Linear regression is used for the task of predicting an attribute, called as the *response variable*, of a dataset given the rest of the attributes, called the *predictor variables*, of the same dataset. The linear regression models assume the response variable is a linear combination of the predictor variables. This is called the linear function hypothesis and is the basis of any general linear model [16]. Any linear regression algorithm finds the parameters of the linear combination for a given *training dataset*. The optimal set of parameters minimizes a *loss function*, such as root mean square error, for the *training dataset* (Section 3.2). Following the *training step*, these models are used for predicting the unknown response variable for the known predictor variables of the *testing dataset*. Linear regression models also assume that the predictor variables are statistically independent to each other [16]. In contrary, real world datasets often contain attributes that are correlated to each other. Such datasets violate the independence assumption, and yield unstable and overfitted solutions of linear regression. regularisation terms are added to the loss function of linear regression to overcome these problems (Section 3.3). Regularisation terms are weighted function of the parameters of the linear regression. *LASSO* [23], *ridge* [13], and *elastic net* [27] are three variants of regularised linear regression. LASSO, ridge and elastic net add regularisation terms proportional to L_1 norm, L_2 norm and a convex combination of L_1 and L_2 norms respectively.

Differential privacy inducing mechanisms, such as the *functional mechanism* [26], are studied and developed for the linear regression models. The functional mechanism [26] adds a calculated amount of noise in the loss function of the linear regression model (Section 4). This mechanism provides a formal privacy guarantee for the linear regression model trained using the noisy loss function. In order to establish the privacy guarantee it leverages the probabilistic framework of ϵ-differential privacy [8]. [26] also instantiate the functional mechanism for the logistic regression.

In this paper, we empirically and comparatively evaluate (Section 5) the functional mechanism for linear regression, and three of its regularised variants, namely, LASSO, ridge, and elastic net. We comparatively evaluate the performance of these four mechanisms and their differentially private variants on two datasets with different correlations and sparsity. We observe that the functional mechanism applied to the regularised linear regression yields similar performance

results and that the private linear regression models perform worse than the non-private linear regression models. We compare the effectiveness of the functional mechanism with an *input perturbation mechanism* [15]. For a given privacy level, ϵ, we empirically show that the functional mechanism is more effective than [15]. We extend the analysis in [26] to empirically study the robustness of the functional mechanism. The key observation in our experiments is that all the private linear regression models are unstable. Our analysis (Section 5.4) shows that the reason for such an instability is inherent to the functional mechanism. The functional mechanism does not necessarily preserve the convexity properties of the loss function after the addition of noise. This potential non-convexity causes instability in the optimisation. In reference to these experimental evidences, we conclude by puting forth (Section 6) the need of designing a differentially private mechanism that produces a convex noisy loss function in order to provide both stable and private output for linear regression models.

2 Related Work

Linear regression [16] is a fundamental yet a widely used [2, 4, 14, 21] machine learning model. Variants of linear regression, Ridge [13] and LASSO [23], are used to reduce correlation in the data features and to avoid overfitting. Elastic net [27] regression uses convex combination of regularisation terms that are used in Ridge and LASSO. For a detailed presentation and discussion of regularisation and regression analysis, interested readers can refer to [16].

Differential Privacy [8] is a probabilistic framework that quantifies the privacy of a randomized function or algorithm. Existing deterministic machine learning models can be randomized by introducing calibrated random noise. The resultant randomized *mechanism* can be shown to satisfy constraints of differential privacy. Dwork et. al. propose the Laplace mechanism [7], which perturbs the output of a machine learning model by explicitly adding scaled random noise from the Laplace distribution. The Gaussian mechanism [8] and the K-norm mechanism [12] are differentially private mechanisms that are also based on the idea of output perturbation with noise from different distributions. Lei [15] proposes differentially private M-estimators, which perturbs the histogram of input data using a scaled noise and further uses the noisy histogram to train the models. Zhang et. al. [26] propose a differentially private *functional mechanism* that adds a properly scaled Laplace noise to the coefficients of loss function in the polynomial basis. Hall et.al. [10] also propose a differentially private *functional mechanism* that adds a properly scaled noise drawn from the Gaussian process to the coefficients of loss function in the kernel basis.

Zhang et. al. instantiate their functional mechanism on linear regression and logistic regression. In order to alleviate the non-convexity caused in the loss function due to addition of random noise, they use ridge regularised linear and logistic regressions. Yu et. al. [25] achieve differential privacy in the elastic net logistic regression by controlling the coefficient of regularisation term. The regularisation term in their proposal is inversely proportional to the number of

datapoints. It causes reduction in regularisation as the number of datapoints increases. Therefore, their proposed mechanism is not applicable for large datasets. Talwar et. al. [22] propose a differentially private variant of Frank-Wolfie optimisation algorithm to perform LASSO regression. This method adds noise in the optimisation algorithm instead of adding it to the objective function.

In this work, we empirically evaluate the effectiveness of using the functional mechanism [26] for linear regression and regularised regression with ridge, LASSO and elastic net regulariser. Unlike the output perturbation mechanisms [7, 8, 12], which study differential privacy under the release of the outputs of the models, we evaluate the differential privacy of linear regression under the release of parameters of the model [15, 26].

3 Background

Let, $\mathcal{D}$ denotes a universe of d-dimensional real-valued datapoints and the corresponding real-valued responses. An element from this universe can be represented by a pair $D = (X, y)$ where $X \in \mathbb{R}^{n \times d}$ denotes a data matrix whose each row corresponds to a d-dimensional datapoint, x_i and $y \in \mathbb{R}^n$ denotes the response vector in one-to-one correspondence to n datapoints. We use $\|\cdot\|_p$ to represent L_p norm of a vector. A dataset D is split into two disjoint sets of datapoints: training dataset, T and validation dataset V.

3.1 Differential Privacy

Let, D' denotes a dataset which differs from the dataset D by one datapoint. Such a dataset D' is said to be a dataset *neighbouring* to the dataset D. Differential privacy uses the notion of *neighbouring* datasets to quantify the privacy of a randomized function, termed as a mechanism $\mathcal{M}$, run on those datasets.

Definition 1. *[8] A randomized algorithm $\mathcal{M}$ with domain $\mathcal{D}$ is ϵ-differentially private if for all $S \in Range(\mathcal{M})$ and $D, D' \in \mathcal{D}$ such that D and D': are neighbouring datasets*

$$\log\left(\left|\frac{Pr(\mathcal{M}(D) \in S)}{Pr(\mathcal{M}(D') \in S)}\right|\right) \leq \epsilon$$

Definition 1 requires $\mathcal{M}$ to be a randomized function. Researchers devise different mechanisms that randomize deterministic function by adding calibrated random noise to either the outputs [7] or *the function* [10, 26] itself to make it *randomized*.

Differential privacy is related to *smoothness* of a function. If a function is smooth, the output of the function does not drastically change when two closely situated inputs are given to the function. A similar interpretation can be made out from the Definition 1. Differential privacy bounds the probability of generating same output by a randomized mechanism $\mathcal{M}$ when two neighbouring datasets are given as inputs. In order to quantify differential privacy, we need a measure to calculate *smoothness* of a function.

Definition 2. *The L_1-sensitivity of a function $f : \mathcal{D} \to \mathbb{R}^k$ is defined as:*

$$\Delta_f = \max_{x,y \in \mathcal{D}, x \sim y} \|f(x) - f(y)\|_1$$

L_1-sensitivity captures the maximum deviation in the output of a query f when two neighbouring datasets are given as input.

Laplace mechanism [7] is one of the widely used mechanism that achieves differential privacy by adding random noise from a properly scaled Laplace distribution.

Definition 3. *The Laplace Distribution with mean zero and scale $b > 0$ is a probability distribution with probability density function:*

$$Lap(x|b) = \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right)$$

where $x \in \mathbb{R}$. We write $Lap(b)$ to denote a random variable $X \sim Lap(x|b)$

Definition 4. *Given any function $f : \mathcal{D} \to \mathbb{R}^k$, the Laplace Mechanism is define as:*

$$\mathcal{M}_\epsilon(x, f) = f(x) + (L_1, ..., L_k)$$

for any $x \in \mathcal{D}$ and $L_i s$ are random variables drawn from $Lap(\frac{\Delta_f}{\epsilon})$.

Theorem 1. *[8] The Laplace mechanism satisfies ϵ-differential privacy.*

3.2 Linear regression

Regression is a kind of predictive machine learning model that learns function $f : \mathbb{R}^d \to \mathbb{R}$ parameterized by a set of parameters of θ. Let, $l_\theta : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ denotes loss function that quantifies the loss in prediction for a given datapoint x_i and the corresponding response y_i due to parameter θ. Learning comprises of the training step wherein one estimates parameters θ which minimize the loss over the training dataset. Mathematically,

$$\theta^* = \arg\min_\theta \sum_{(x_i, y_i) \in T} l_\theta(x_i, y_i) \tag{1}$$

Linear regression uses a linear hypothesis to map predictor variable to the corresponding response. In matrix notation, linear regression is parameterized by $\theta \in \mathbb{R}^d$ such that $X\theta = y$. In order to find the optimal value of θ, training step in linear regression minimizes *mean squared loss* over the training data as defined in Equation 2.

$$\theta^* = \arg\min_\theta \sum_{(x_i, y_i) \in T} (x_i^t \theta - y_i)^2 = \arg\min_\theta (X\theta - y)^2 \tag{2}$$

Linear regression problem can be analytically solved by differentiating Equation 2 and equating it to zero. The analytical solution is:

$$\theta^* = (X^t X)^{-1} X^t y \tag{3}$$

3.3 Regularised Linear Regression

Properties of the coefficient of quadratic term in Equation 2, *viz.* X^tX, determine the convexity of the optimisation problem[1]. If the matrix X^tX does not possess a full rank due to correlation among different data dimensions, it does not remain a positive-definite matrix and induces non-convexity in the corresponding optimisation problem. Non-convex optimisation problems do not guarantee a unique global minimum.

In order to alleviate these irregularities, a constant term is added to the diagonal elements of the matrix X^tX which forcefully makes it a full rank matrix. Under such a transformation, the linear regression optimisation problem transforms into the optimisation problem for *ridge regression* [13] as stated in Equation 4. Closed form solution for the ridge regression is presented in Equation 5.

$$\theta^* = \arg\min_{\theta}(X\theta - Y)^2 + \lambda\|\theta\|_2^2 \tag{4}$$

$$\theta^* = (X^tX + \lambda\mathbb{I}_d)^{-1}X^ty \tag{5}$$

Addition of the regularisation term that is proportional to L_1 norm of parameter yields a variant of linear regression called as *LASSO regression* [23]. LASSO regression is used for obtaining *sparse* solution for the regression problem. Since a sparse solution contains many zeros, only the features that have non-zero coefficients contribute towards generating the response. Due to this property, LASSO regression is also used for variable selection. Unlike linear regression and ridge regression, we do not have a closed form solution for LASSO regression due to piecewise differentiability of L_1 norm.

4 Functional Mechanism for regularised Linear Regression

Unlike output perturbation mechanisms, Functional mechanism [26] introduces random noise in the loss function of a machine learning algorithm. optimisation of such a noisy loss function leads to the parameters which are different than true optimal parameters. In this way, we indirectly get noisy outputs from the machine learning model without explicitly adding noise to the outputs themselves. In this section, we elucidate the details related to the functional mechanism.

4.1 Sensitivity calculation

By Stone-Weierstrass theorem, any infinitely differentiable function can be expanded in terms of its polynomial basis for some $J \in [0, \infty]$ as given in Equation 6

[1] X^t denotes the transpose of X.

where Φ_j denotes the set of polynomials with degree j and $\lambda_{t\phi}$ denote respective coefficients.

$$f(t,\omega) = \sum_{j=0}^{J} \sum_{\phi \in \Phi_j} \lambda_{t\phi} \phi(\omega) \tag{6}$$

For a given machine learning model, the loss function l_θ can be expanded in the polynomial basis as a function of parameter θ as given in Equation 7 where $t = (x, y)$ denotes a datapoint in training dataset T.

$$l(T,\theta) = \sum_{t \in T} \sum_{j=0}^{J} \sum_{\phi \in \Phi_j} \lambda_{t\phi} \phi(\theta) \tag{7}$$

Lemma 1. *[26] L_1 sensitivity of the loss function of a machine learning model is given by:*

$$\begin{aligned} \Delta_l &= \max_{D, D' \in \mathcal{D}, D \sim D'} \| l(D,\theta) - l(D'\theta) \|_1 \\ &= 2 \max_t \sum_{j=1}^{J} \sum_{\phi \in \Phi_j} \| \lambda_{t\phi} \|_1 \end{aligned}$$

where $t = (x, y)$ denotes a datapoint in D or D'.

4.2 Functional Mechanism

Using Theorem 1 and the Laplace mechanism, Zhang et. al. [26] devise an algorithm that adds noise to the loss function. For the sake of completeness and notational consistency, we present the algorithm in Algorithm 1.

Algorithm 1 Functional Mechanism [26](Training Dataset T, loss function l_θ, Privacy Parameter ϵ)

1: $\Delta_l = 2 \max_t \sum_{j=1}^{J} \sum_{\phi \in \Phi_j} \| \lambda_{t\phi} \|_1$
2: **for** each $j \in [0, 1, ..., J]$ **do**
3: **for** each $\phi \in \Phi_j$ **do**
4: $\lambda_\phi \leftarrow \sum_{t \in T} \lambda_{t\phi} + \text{Lap}(\frac{\Delta_l}{\epsilon})$
5: **end for**
6: **end for**
7: Let, the noisy loss function $l'(T,\theta) \leftarrow \sum_{j=0}^{J} \sum_{\phi \in \Phi_j} \lambda_\phi \phi(\theta)$
8: Compute $\theta^* = \arg\min_\theta l'(T,\theta)$
9: **return** θ^*

Theorem 2. *Algorithm 1 is satisfies ϵ-differential privacy.*

Proof. Proof is available in [26].

4.3 Case: Elastic Net Regression

Now we apply the idea of functional mechanism to a specific case of machine learning model, namely linear regression. We know that linear regression models use *squared error* as the loss function, which is minimized to find the parameters. We expand the loss function, stated in Equation 2, as a polynomial in the parameters, θ, for a given data matrix X and the corresponding response vector y.

$$l(T, \theta) = \theta^t(X^tX)\theta - 2\theta^t(X^ty) + y^Ty \tag{8}$$

Elastic net regression [27] adds the regularisation term which is a convex combination of L_1 regularisation term and L_2 regularisation term. The optimisation problem for elastic net regression with functional mechanism is given in Equation 9. $l'(T, \theta)$ denotes the noisy loss function obtained by applying Algorithm 1 on the loss function for linear regression, as stated in Equation 8.

$$\theta^* = \arg\min_{\theta} l'(T, \theta) + \lambda(\alpha\|\theta\|_2^2 + (1-\alpha)\|\theta\|_1) \tag{9}$$

If we put $\alpha = 1$ in the Equation 9, the objective function reduces to the objective function of functional mechanism with ridge regression. If we put $\alpha = 0$ in the Equation 9, the objective function reduces to the objective function of functional mechanism LASSO regression. Therefore, elastic net regularisation stands as a bridge between these two variants.

As stated in the Section 3, a regularised linear regression incorporates a regularisation term, which is a term proportional to norm of the parameters, in the objective function. The sensitivity of a regularised objective function remains same as the sensitivity of an non-regularised objective function due to independence of the regularisation term on the dataset. Therefore, calculations in Algorithm 1 do not change under the regularisation. We use Lemma 1 to calculate sensitivity of the loss function stated in Equation 8.

Lemma 2. *Assuming that all features of datapoints are normalized such that each of the feature value lies in* $[-1, 1]$*,* L_1 *sensitivity of the loss function in Equation 8 is given by:*

$$\Delta_l = 2(d^2 + 2d + 1)$$

5 Empirical Performance Evaluation

We comparatively and empirically evaluate functional mechanism for regularised linear regressions: namely ridge, LASSO, and elastic net. We present the result analysis in this section.

5.1 Datasets

We conduct experiments on a microdata sample of US Census in 2000 provided by IPUMS International [1]. The **census dataset** consists of 1% sample of the

original census data. It spans over 1.23 million households with records of 2.8 million people. It has several attributes of which not every single attribute is reported by all of the people. In order to avoid the discrepancies in the data, we consider $316,276$ records of the heads of households in our dataset. Each record has 9 attributes, namely, *Age, Gender, Race, Marital Status, Family Size, Education, Employment Status, House type, Income.* Regression analysis is performed using *Income* as the response variable and the rest of the attributes as predictor variables.

Correlation among different attributes of a dataset adds training bias in the effectiveness of the model. In order to derive unbiased inferences from the results, we use an another dataset: **wine quality testing dataset** [5]. The dataset comprises 4898 records of white wine samples wherein each record has 12 attributes, namely, *Fixed Acidity, Volatile Acidity, Citric Acid, Residual Sugar, Chlorides, Free Sulfur Dioxide, Total Sulfur Dioxide, Density, pH, Sulphates, Alcohol, and Quality.* Regression analysis is performed using *Quality* as the response variable and the rest of the attributes as predictor variables.

5.2 Experimental Setup

All programs are run on Linux machine with 12-core 3.60GHz Intel® Core i7™processor with 64GB memory. Python® 2.7.6 is used as the scripting language. We use SCS [17] solver, that is available in CVXPY [6] package, to find the solution for piecewise differentiable objective functions of LASSO and elastic net regression.

We report the results as the aggregates over 50 experimental runs. For every experimental run, we randomly hold out 20% of the data for testing and use the rest 80% of the data for training regression models. We normalize each of these features such that their values lie in $[-1, 1]$. We use *root mean squared error (RMSE)* as the metric to comparatively evaluate effectiveness. We report RMSE for a trained model calculated on the validation dataset as given in the Equation 10. For given value of ϵ, the model with smallest value of RMSE is the most effective model.

$$\text{RMSE}_\theta = \sqrt{\frac{\sum_{(x_i,y_i)\in V}(x_i^t\theta - y_i)^2}{|V|}} \tag{10}$$

We comparatively evaluate eight regression problems: linear regression (LR), ridge regression (RG), LASSO regression (LS), elastic net regression (EN), and their private versions, the ones that are obtained applying functional mechanism to each of these four. For brevity, we call the regression model obtained using the functional mechanism as *functional regression.* For every regularised regression model, we set the regularisation coefficient, λ, by performing cross-validation on the respective non-private versions of regularised regression and choosing the value that results in smallest in testing error. We use the same regularisation coefficient for the private version.

5.3 Results

Figure 1 shows the boxplot of functional elastic net regression for different values of ϵ's for the census dataset. We note the presence of a large number of outliers in the result. We observe similar results for the rest of the functional regressions. Use of *mean* as an aggregate affects the analysis of the results due to sensitivity of mean to outliers in the data. In order to avoid this bias due to the outliers, we choose *median* as an aggregate to report the results. Due to lack of space, we do not present the results with *mean aggregate*.

Comparative evaluation of Functional Mechanism for Regularised Linear Regression. Figure 2 and Figure 3 show the comparative evaluation of the variants of regularised linear regression for the wine quality dataset and the census dataset respectively. In the plot, solid line represents median over 50 experimental runs and the shaded region covers RMSE values that lie between 20^{th} and 80^{th} percentile. Smaller values of ϵ's induce higher noise in the function, which in turn results in higher privacy. Therefore, we observe higher RMSE for smaller values of ϵ's. As the value of ϵ increases, the effectiveness of the functional regression approaches the the effectiveness of the non-private counterpart.

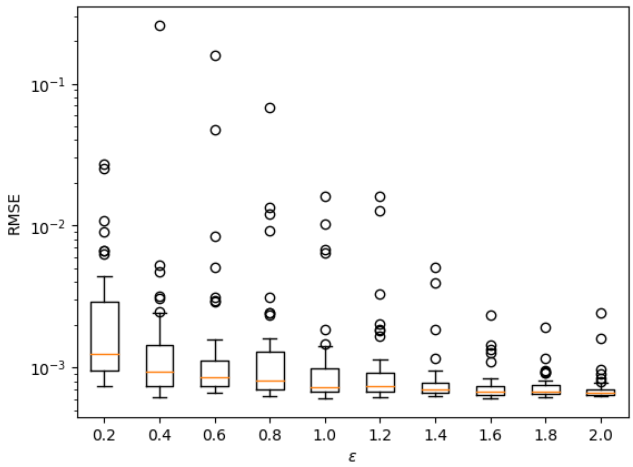


Fig. 1: Boxplot of RMSE of elastic net ridge regression with functional mechanism for different values of ϵ for the census dataset.

In order to understand the magnitude of instability, we plot the variance of different regressions in Figure 4. We observe that the non-private models show vary small amount of variance. As the value of ϵ increases, the amount of noise that is added in the loss function reduces. Reduction in the noise results in lower RMSEs and lower variance in RMSE. We do not observe the same trend in functional linear regression. In the absence of regularisation, noisy loss function

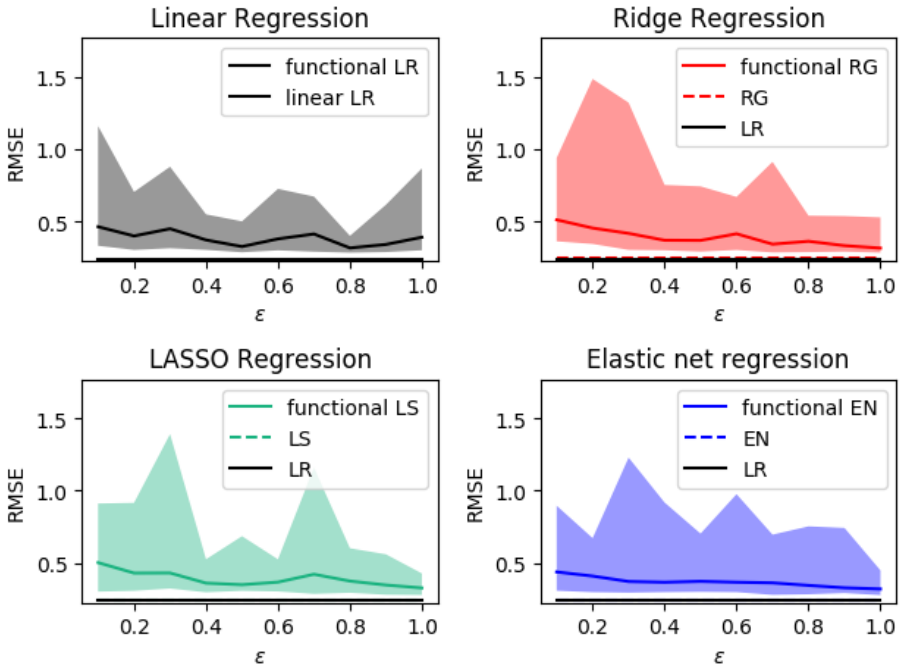


Fig. 2: Comparison of different regressions for the wine quality testing dataset with *median* the as aggregate

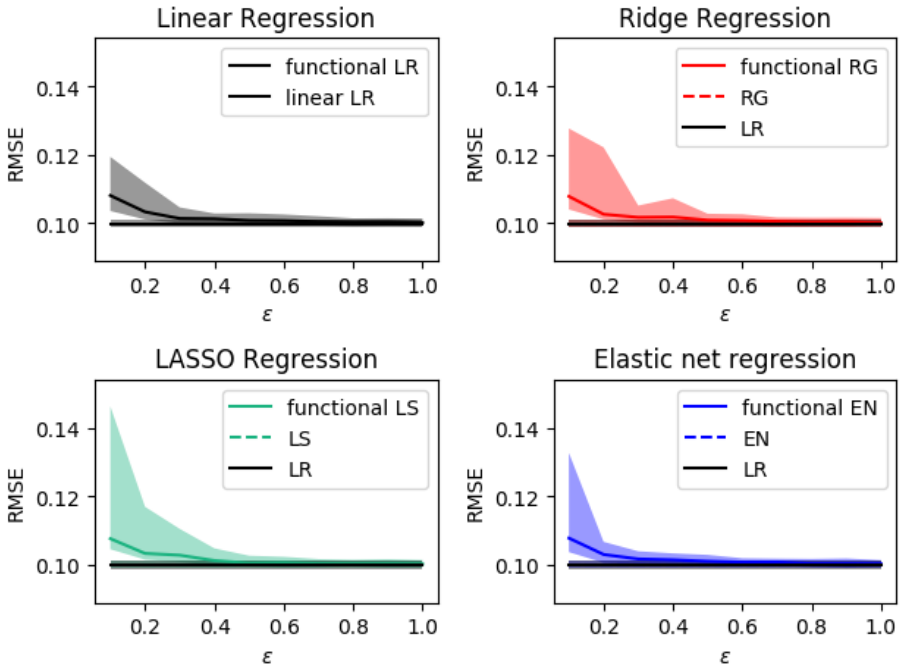
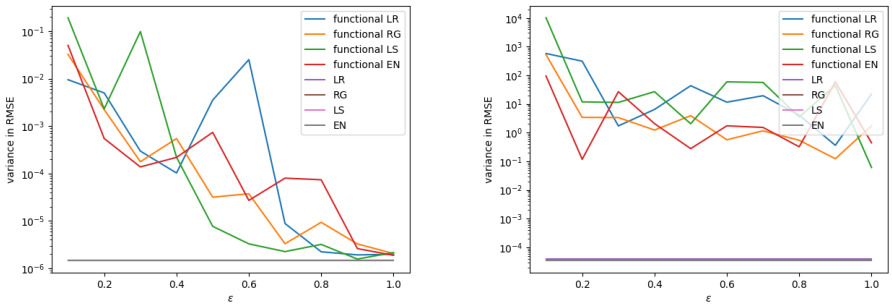


Fig. 3: Comparison of different regressions for the census dataset with *median* as the aggregate



(a) Census Dataset (b) Wine quality testing dataset

Fig. 4: Variance in RMSE for varying values of ϵ's for different regressions.

of linear regression becomes highly non-convex. Therefore, the results show high instability.

Comparative evaluation of Functional Mechanism and Input Perturbation Mechanism. We, also, compare effectiveness of the functional mechanism with an *input perturbation mechanism*, differentially private M-estimators (DPME) [15]. Discretisation of large number of variables leads to a large discrete space that causes prohibitive computation cost. Due to concentration of data around subsets of features, a large discrete space also leads to sparse histograms [15]. In order to alleviate the sparsity, we follow [15] and evaluate the performance on a simpler regression model. We show the comparative study on the census dataset where we predict *Income* of a person using *Age, Gender, Race* and *Education Status* as the predictors [2]. The results are presented in Figure 5. Solid lines represent the *mean* RMSE over 50 experimental runs. For a given value of ϵ, we observe that the functional mechanism provides lower RMSE for all regularised linear regressions. M-estimator is a robust statistic [11]. Therefore, we observe the stability in the performance of DPME as compared to the results presented in Figure 3 and Figure 2.

5.4 Result Analysis

One observation that is common in all the empirical evaluations is the instability in the results. In order to achieve privacy, we, indeed, need to introduce noise in the output so that it does not reveal the true output of the model.

[2] The wine quality testing dataset comprises of 4898 datapoints over 12 continuous attributes. Due to the small size of the dataset, it generates highly sparse histograms in the large space. Therefore, for the sake of interpretability, we present the results for the census dataset.

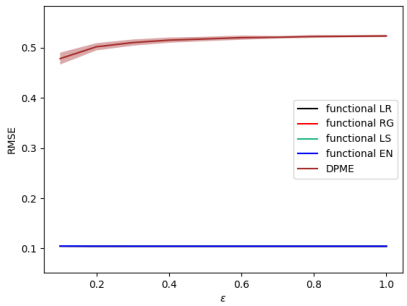


Fig. 5: Comparative evaluation of DPME [15] and functional mechanism for the census dataset.

But excessive noise deteriorates the effectiveness, and hence the utility, of the machine learning model. We find that the reasons for this instability are rooted in the functional mechanism it self.

Symmetric Noise. The coefficient of the quadratic term in Equation 8, X^tX, is a symmetric matrix. It loses the symmetric property after adding random noise from the Laplace distribution. A standard way to make a given matrix A symmetric is to use $(A + A^t) * 0.5$. This way of symmetrization of noisy X^tX indirectly incurs addition two Laplace random variables. Addition of two Laplace random variable does not follow Laplace distribution. Therefore, in order to maintain the integrity of the functional mechanism, we can not make X^tX symmetric in the conventional way.

Convexity of the objective function. Linear regression works on the assumption that the features of data are independent of each other. Independence among the features makes X^tX a positive definite matrix. Positive definite matrices make the optimisation convex and guarantees optimality of the solution. Noisy loss function fails to guarantee convexity of the objective problem, and hence the optimality of the solution. A similar observation is made by Lei [15] while perturbing the histograms of input data by adding the calibrated noise. In order to make the objective function convex, Zhang et. al. [26] calculate the spectral decomposition of X^tX and consider the projection of parameters onto the eigenspace spanned by eigenvectors with positive eigenvalues. They do not provide any analytical justification which guarantees differential privacy after pruning the non-positive eigenspace.

Differential privacy of the optimisation algorithm. Functional mechanism proves that the loss function generated by any two neighbouring datasets satisfies differential privacy. Composition of a differentially private function with a *deterministic* function, called as *post-processing* [8], remains differentially private. An optimisation problem solver calculates an approximate solution when the objective function is not convex. Therefore, differential privacy of a loss function is not preserved by the optimisation algorithm itself.

6 Conclusion and Future Works

In this work, we use functional mechanism that presents differentially private linear regression models. We empirically and comparatively evaluate the effectiveness of the private and non-private versions of linear regression, LASSO, ridge and elastic net on the census and the wine quality datasets. For a given privacy level, ϵ, we observe that the functional mechanism is more effective than DPME [15] for the regularized linear regression. We extend the analysis in [26] to empirically study the robustness of the functional mechanism. We invariably observe that the private versions are less effective than their non-private counterparts. The key observation from these experiments is that all these private regularised regression methods are equally unstable, and private linear regression is comparatively more unstable.

We analyze the loss of symmetry of the covariance matrix and the non-convexity of the loss function after adding the noise as the principal reasons of this instability. Thus, additional structure of noise and loss function are required to be known in order to keep the loss function of linear regression convex and stable even after adding the noise. This opens up the need of designing a differential privacy mechanism that would retain these properties for private linear regression mechanisms. This is an interesting theoretical question along with practical applications. We are now looking into development of such mechanism. Development of such mechanism would also be relevant for all the machine learning methods, such as empirical risk minimization [3], that optimise a loss function computed on a dataset.

References

1. Minnesota population center. integrated public use microdata series international: Version 5.0. https://international.ipums.org. (2009)
2. Aalen, O.O.: Further results on the non-parametric linear regression model in survival analysis. Statistics in medicine 12(17), 1569–1588 (1993)
3. Chaudhuri, K., Monteleoni, C., Sarwate, A.D.: Differentially private empirical risk minimization. Journal of Machine Learning Research 12(Mar), 1069–1109 (2011)
4. Chi, G., Voss, P.: Migration decision-making: a hierarchical regression approach. Journal of Regional Analysis and Policy 35(2) (2005)
5. Cortez, P., Cerdeira, A., Almeida, F., Matos, T., Reis, J.: Modeling wine preferences by data mining from physicochemical properties. Decision Support Systems 47(4), 547–553 (2009)
6. Diamond, S., Boyd, S.: CVXPY: A Python-embedded modeling language for convex optimization. Journal of Machine Learning Research 17(83), 1–5 (2016)
7. Dwork, C., McSherry, F., Nissim, K., Smith, A.: Calibrating noise to sensitivity in private data analysis. In: Theory of Cryptography Conference. pp. 265–284. Springer (2006)
8. Dwork, C., Roth, A., et al.: The algorithmic foundations of differential privacy. Foundations and Trends® in Theoretical Computer Science 9(3–4), 211–407 (2014)

9. Goodfeloow, I., Papernot, N.: Is attacking machine learning easier than defending it? (2017), http://www.cleverhans.io/security/privacy/ml/2017/02/15/why-attacking-machine-learning-is-easier-than-defending-it.html
10. Hall, R., Rinaldo, A., Wasserman, L.: Differential privacy for functions and functional data. Journal of Machine Learning Research (JMLR) 14(Feb), 703–727 (2013)
11. Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A.: Robust statistics: the approach based on influence functions, vol. 196. John Wiley & Sons (2011)
12. Hardt, M., Talwar, K.: On the geometry of differential privacy. In: Proceedings of the forty-second ACM Symposium on Theory of Computing (STOC). pp. 705–714. ACM (2010)
13. Hoerl, A.E., Kennard, R.W.: Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12(1), 55–67 (1970)
14. Intriligator, M.D.: Econometric models, techniques, and applications. Tech. rep., Prentice-Hall Englewood Cliffs, NJ (1978)
15. Lei, J.: Differentially private m-estimators. In: Advances in Neural Information Processing Systems. pp. 361–369 (2011)
16. Murphy, K.P.: Machine Learning: A Probabilistic Perspective. The MIT Press (2012)
17. O'Donoghue, B., Chu, E., Parikh, N., Boyd, S.: Conic optimization via operator splitting and homogeneous self-dual embedding. Journal of Optimization Theory and Applications 169(3), 1042–1068 (June 2016), http://stanford.edu/ boyd/papers/scs.html
18. Ribeiro, M., Grolinger, K., Capretz, M.A.: Mlaas: Machine learning as a service. In: Machine Learning and Applications (ICMLA), 2015 IEEE 14th International Conference on. pp. 896–902. IEEE (2015)
19. Shokri, R., Stronati, M., Song, C., Shmatikov, V.: Membership inference attacks against machine learning models. In: Security and Privacy (SP), 2017 IEEE Symposium on. pp. 3–18. IEEE (2017)
20. Song, C., Ristenpart, T., Shmatikov, V.: Machine learning models that remember too much. In: Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. pp. 587–601. ACM (2017)
21. Strobl, C., Malley, J., Tutz, G.: An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychological methods 14(4), 323 (2009)
22. Talwar, K., Thakurta, A.G., Zhang, L.: Nearly optimal private lasso. In: Advances in Neural Information Processing Systems (NIPS). pp. 3025–3033 (2015)
23. Tibshirani, R.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) pp. 267–288 (1996)
24. Tramèr, F., Zhang, F., Juels, A., Reiter, M.K., Ristenpart, T.: Stealing machine learning models via prediction apis. In: USENIX Security Symposium. pp. 601–618 (2016)
25. Yu, F., Rybar, M., Uhler, C., Fienberg, S.E.: Differentially-private logistic regression for detecting multiple-snp association in gwas databases. In: International Conference on Privacy in Statistical Databases. pp. 170–184. Springer (2014)
26. Zhang, J., Zhang, Z., Xiao, X., Yang, Y., Winslett, M.: Functional mechanism: regression analysis under differential privacy. Proceedings of the VLDB Endowment 5(11), 1364–1375 (2012)
27. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67(2), 301–320 (2005)