(α, β)-Differential Privacy in Reproducing Kernel Hilbert Spaces

Rob Hall, Alessandro Rinaldo, Larry Wasserman

Abstract

Differential privacy is a framework for releasing summaries of a dataset in a way that makes rigorous the notion of privacy afforded to the data elements (e.g., when they are measurements of individuals). Previous work has focused on the release of finite dimensional statistics where privacy is achieved via randomization. We extend these methods to allow the release of infinite dimensional quantities in an RKHS, which leads to a new technique for the privacy-preserving release of functions such as kernel machines.

1 Differential Privacy

Suppose we have database $X \in \mathcal{X}^n$ and want to release a summary (e.g., a statistical analysis) without compromising the privacy of individuals in the database. One framework for defining privacy rigorously in such problems is *Differential privacy* [3, 5]. We characterize a method for the data release by a family of probability distributions $\{P_X : X \in \mathcal{X}^n\}$ over $(\Omega, \mathcal{A})$, which specify the probabilities of the output ω for each possible input database X. We use the relation $X \sim X'$ to mean that X, X' differ in one element (that is, there exists a permutation of X' having hamming distance one from X). A randomized algorithm is called (α, β)-Differentially Private if, for all $X \sim X' \in \mathcal{X}^n$ we have:

$$P_X(A) \leq e^{\alpha} P_{X'}(A) + \beta, \quad \forall A \in \mathcal{A}. \tag{1}$$

In essence this condition ensures that the output distribution does not change much when the input data changes by one element, the implication being that it is very hard to distinguish between input databases on the basis of the output (see [7]). The parameters α, β control the difficulty of inference about the input, and are used to tradeoff the error due to noise against the strength of the privacy guarantee. Algorithms that fulfil differential privacy have been developed for classification, logistic regression and many other learning tasks. See, for example, [2,6], and references therein. In all these cases, the output are taken to be real vectors (e.g., regression coefficients). Techniques for the release of functions have so far been limited to finite dimensional projections (see e.g., [6, 7]), whereas we give a technique which releases an infinite dimensional quantity. We also make use of the RKHS structure, which makes the techniques more straightforward.

Finally we recall that the archetypal method for the release of a real vector $v_X = v(X) \in \mathbb{R}^d$ is to add Gaussian noise proportional to the "global sensitivity" of that function (see e.g., [4]).

Proposition 1.1. *For some fixed covariance matrix Σ, whenever*

$$\sup_{X \sim X'} \left\| \Sigma^{-1/2} (v_X - v_{X'}) \right\|_2 \leq \Delta, \tag{2}$$

then it satisfies (α, β)-DP to take each P_X to be the multivariate normal distribution having mean v_X, and covariance matrix $c(\alpha, \beta, \Delta)\Sigma$ where we define $c(\alpha, \beta, \Delta) = 2 \log \frac{2}{\beta} \Delta^2 / \alpha^2$.

2 Differential Privacy in an RKHS

Denote by $\mathcal{H}(K)$ the RKHS over $\mathbb{R}^T$ (we restrict attention to the case of $T = [0, 1]^d$) with the reproducing kernel K. When the goal is to release a function $f_X = f(X, \cdot)$ which depends on the data X, differential privacy is achieved by the addition of an appropriate Gaussian process. We first give the analog of proposition 1.1 then give a proof sketch.

Proposition 2.1. *For a family of functions $\{f_X : X \in \mathcal{X}^n\} \subseteq \mathcal{H}(K)$ which satisfies*

$$\sup_{X \sim X'} \|f_X - f_{X'}\|_{\mathcal{H}(K)} \leq \Delta_{\mathcal{H}(K)}, \tag{3}$$

then it satisfies (α, β)-DP to take each P_X to be the Gaussian process measure having mean function f_X and covariance function $c(\alpha, \beta, \Delta_{\mathcal{H}(K)})K$, with c defined as before.

In sketching the proof, we use the spectral representation of the RKHS, namely we write the kernel as $K(x, y) = \sum_{i \geq 1} \lambda_i \phi_i(x) \phi_i(y)$, where the ϕs form an orthonormal basis under the RKHS inner product which we denote

$\langle \cdot, \cdot \rangle_{\mathcal{H}(K)}$. Likewise a function $f \in \mathcal{H}(K)$ may be uniquely written in terms of coordinates in the same basis

$$f = \sum_{i \geq 1} \eta_i(f)\phi_i(\cdot), \quad \eta_i(f) = \lambda_i \langle f, \phi_i \rangle_{\mathcal{H}(K)}.$$

Proof of proposition 2.1. Consider the Karhunen-Loève expansion of the suggested Gaussian process as $\sum_{i \geq 1} Z_i \phi_i$, where each Z_i is normal with mean $\eta_i(f_X)$ and variance $c(\alpha, \beta, \Delta_{\mathcal{H}(K)})\lambda_i$. If we were to truncate this expansion after m terms, we could interpret the value as a finite dimensional Gaussian vector. Denoting by Σ the diagonal matrix with elements given by λ_i, we note that (3) implies that $\Delta_{\mathcal{H}(K)}$ is an upper bound as required in (2) where $v_{X,i} = \eta_i(f_X)$, and so the truncation admits the differential privacy per proposition 2.1. To demonstrate privacy of the full expansion, we we use a limiting argument. Denote an arbitrary measurable set of the infinite sequence of coefficients by $A = \bigcap_{i \geq 1} A_i$ where A_i is a set of infinite sequences in $\mathbb{R}$ in which the i^{th} is restricted to lie in some measurable set. Since A is the limit of a decreasing sequence of sets $B_m = \bigcap_{i=1}^{m} A_i$, we have $P_X(A) = P_X(\lim_{m \to \infty} B_m) = \lim_{m \to \infty} P_X(B_m)$, and since differential privacy holds for each finite m, a simple limiting argument leads to the requisite privacy of the Gaussian process. □

We remark that under the restriction that T be compact, the error incurred due to privacy, when measured in mean $\mathcal{L}_2$ error, is given by the expectation of the square norm of the Gaussian process which is $\lambda(T)c(\alpha, \beta, \Delta) \sup_{x \in T} K(x, x)$, where we use λ to mean the Lebesgue measure (namely $\lambda(T) = 1$ for $T = [0, 1]^d$).

3 Examples

A first example is kernel density estimation, where $f_X(x) = \frac{1}{n} \sum_{i=1}^{n} K(x_i, x)$, in which K is e.g., the isotropic Gaussian kernel having standard deviation σ (also called the "bandwidth" in the context of density estimation). Since these functions are in the generating set of the RKHS it is easy to find that for $X \sim X'$,

$$\|f_X - f_{X'}\|_{\mathcal{H}(K)} \leq \frac{2}{n} \sup_x \|K(x, \cdot)\|_{\mathcal{H}(K)} = \frac{2}{n(\sqrt{2\pi}\sigma)^{d/2}}.$$

Therefore we find that our proposed method adds error on the order of $n^{-2}\sigma^{-2d}$, which is at a rate far faster than that of the sampling error of $\sigma^4 + n^{-1}\sigma^{-d}$, and permits the optimal choice of $\sigma = n^{-1/(1+4d)}$.

Another example is the release of a regularized loss minimization in an RKHS (namely a kernel machine). When

$$f_X = \arg \min_{f \in \mathcal{H}(K)} \frac{1}{n} \sum_{i=1}^{n} \ell(f, x_i) + \lambda \|f\|^2_{\mathcal{H}(K)},$$

for some loss function ℓ and where λ is a parameter used to control the regularization, then following [1] we find a valid upper bound on the "sensitivity" to be $\frac{L}{n\lambda} \sqrt{\sup_{x \in T} K(x, x)}$, in which L is the Lipschitz constant on ℓ (namely $L = 1$ for hinge loss). Thus once again the error due to the noise addition does not interfere with the optimal rate.

Acknowledgments

This research was partially supported by Army contract DAAD19-02-1-3-0389 to Cylab, and NSF Grants BCS0941518 and SES1130706 to the Department of Statistics, both at Carnegie Mellon University.

References

[1] O. Bousquet and A. Elisseeff. Stability and generalization. *Journal of Machine Learning Research*, 2:499–526, 2002.

[2] K. Chaudhuri and C. Monteleoni. Privacy preserving logistic regression. *NIPS 2008*, 2008.

[3] C. Dwork. Differential privacy. *33rd International Colloquium on Automata, Languages and Programming*, pages 1–12, 2006.

[4] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. *EUROCRYPT*, pages 486–503, 2006.

[5] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. *Proceedings of the 3rd Theory of Cryptography Conference*, pages 265–284, 2006.

[6] B.I.P. Rubinstein, P.L. Bartlett, L. Huang, and N. Taft. Learning in a large function space: Privacy-preserving mechanisms for svm learning. *Journal of Privacy and Confidentiality*, 2010.

[7] L. Wasserman and S. Zhou. A statistical framework for differential privacy. *The Journal of the American Statistical Association*, 105:375–389, 2010.