Theory of Differential Privacy

EN.601.438/638, Fall 2026

Problem Set 3: Privacy Accounting and Optimization

100 points • 15% of course grade • Due: November 4, 23:59

Submission. Submit one PDF on Gradescope. Include your name and list, for each problem, any collaborators and outside resources or tools you used. Unless a problem explicitly asks only for discussion, justify every answer with a proof.
Collaboration. You may discuss definitions, concepts, and high-level ideas with classmates and course staff. You may not view another student's written solution, jointly compose a proof, or reuse text. After a discussion, put away your notes and write your solution independently from scratch.
Generative AI. You may use an AI tool as a peer to clarify definitions, request a new example, or quiz yourself. Do not paste the assignment questions into a tool, request a solution or solution outline, keep the assignment open while using the tool, or submit AI-generated reasoning. Close the tool before writing your solution, and disclose how you used it.
Conventions. All logarithms are natural. Unless stated otherwise, privacy and failure parameters lie in $(0, 1)$. Problem 1 uses add/remove adjacency; Problems 2 and 3 use replace-one adjacency when privacy is discussed. All norms are Euclidean norms unless labeled otherwise.

Facts you may use

- If $\mathcal{W} \subseteq \mathbb{R}^d$ is nonempty, closed, and convex, then its Euclidean projection satisfies

$$\|\Pi_{\mathcal{W}}(u) - v\|_2 \leq \|u - v\|_2 \qquad \text{for every } u \in \mathbb{R}^d \text{ and } v \in \mathcal{W}.$$

- If $Z \sim \mathcal{N}(0, \tau^2 I_d)$, then $\mathbb{E}[Z] = 0$ and $\mathbb{E}[\|Z\|_2^2] = d\tau^2$.
- *For Problem 2(d) only:* a Gaussian mechanism with ℓ_2-sensitivity Δ and noise $\mathcal{N}(0, \tau^2 I_d)$ is $\Delta^2/(2\tau^2)$-zCDP. zCDP parameters add under adaptive composition and are preserved by post-processing.

Problem 1. Amplification by Poisson subsampling [30 points]

In this problem, datasets use *add/remove adjacency*: two finite, record-indexed datasets are neighboring when one can be obtained from the other by adding or removing one record. Let $\mathcal{A}$ be an ε-differentially private mechanism under this adjacency relation, defined on datasets of every finite size. Thus, for every pair of neighboring datasets S, S' and every output event E,

$$\Pr[\mathcal{A}(S) \in E] \le e^{\varepsilon} \Pr[\mathcal{A}(S') \in E].$$

For $q \in [0,1]$, let $\mathsf{Sub}_q(D)$ include each record of D independently with probability q, and define

$$\mathcal{A}_q(D) = \mathcal{A}\big(\mathsf{Sub}_q(D)\big).$$

The subsample itself is not released. Prove that $\mathcal{A}_q$ is ε_q-differentially private under add/remove adjacency, where

$$\boxed{\varepsilon_q = \log\big(1 + q(e^{\varepsilon} - 1)\big).}$$

(a) **Coupling the two subsamples [8 points]** Let $D^+ = D \uplus \{x\}$. Couple the subsampling procedures by making identical inclusion decisions for the records already in D. For an arbitrary output event E, set

$$a = \Pr\big[\mathcal{A}\big(\mathsf{Sub}_q(D)\big) \in E\big], \qquad b = \Pr\big[\mathcal{A}\big(\mathsf{Sub}_q(D) \uplus \{x\}\big) \in E\big],$$

where the probabilities include the subsampling and the internal randomness of $\mathcal{A}$. Show that

$$b \le e^{\varepsilon} a, \qquad a \le e^{\varepsilon} b,$$

and prove the mixture identity

$$\Pr[\mathcal{A}_q(D^+) \in E] = (1-q)a + qb.$$

(b) **The amplified guarantee [17 points]** Use part (a) to prove both privacy directions:

$$\Pr[\mathcal{A}_q(D^+) \in E] \le e^{\varepsilon_q} \Pr[\mathcal{A}_q(D) \in E] \quad \text{and} \quad \Pr[\mathcal{A}_q(D) \in E] \le e^{\varepsilon_q} \Pr[\mathcal{A}_q(D^+) \in E].$$

For the reverse direction, it may be useful first to verify that

$$\big(1 - q + qe^{\varepsilon}\big)\big(1 - q + qe^{-\varepsilon}\big) \ge 1.$$

(c) **Interpretation [5 points]** Show that, for $0 < \varepsilon \le 1$,

$$\frac{q\varepsilon}{e} \le \varepsilon_q \le (e-1)q\varepsilon.$$

Conclude that $\varepsilon_q \approx q\varepsilon$ uniformly for $0 < \varepsilon \le 1$.

Problem 2. Noisy projected gradient descent [40 points]

Let $\mathcal{W} \subseteq \mathbb{R}^d$ be nonempty, closed, and convex, with Euclidean diameter at most R. For a fixed dataset $D = (x_1, \dots, x_n)$, let

$$L_D(w) = \frac{1}{n} \sum_{i=1}^{n} \ell(w; x_i),$$

where every $\ell(\cdot; x)$ is differentiable and convex on $\mathcal{W}$, and

$$\|\nabla_w \ell(w; x)\|_2 \le G \qquad \text{for every } w \in \mathcal{W} \text{ and every } x.$$

Thus $\|\nabla L_D(w)\|_2 \le G$. Assume L_D attains its minimum over $\mathcal{W}$, and let $w^\star \in \arg\min_{w\in\mathcal{W}} L_D(w)$. Starting from a public, data-independent point $w_1 \in \mathcal{W}$, run, for $t = 1, \dots, T$,

$$\widetilde{g}_t = \nabla L_D(w_t) + Z_t, \qquad Z_t \sim \mathcal{N}(0, \tau^2 I_d), \qquad w_{t+1} = \Pi_{\mathcal{W}}(w_t - \eta \widetilde{g}_t),$$

and return

$$\bar{w}_T = \frac{1}{T}\sum_{t=1}^{T} w_t.$$

Here $\eta > 0$. Assume that $Z_1, \dots, Z_T$ are mutually independent and independent of w_1 and all earlier algorithmic randomness. Let $\mathcal{F}_{t-1}$ denote the randomness available immediately before drawing Z_t; in particular, w_t is $\mathcal{F}_{t-1}$-measurable. All expectations below are over the algorithmic randomness, with D held fixed.

(a) **A potential inequality [10 points]** Prove that, for every $u \in \mathcal{W}$,

$$\langle \widetilde{g}_t, w_t - u\rangle \le \frac{\|w_t - u\|_2^2 - \|w_{t+1} - u\|_2^2}{2\eta} + \frac{\eta}{2}\|\widetilde{g}_t\|_2^2.$$

(b) **Conditional expectations [10 points]** Prove the two conditional identities

$$\mathbb{E}[\langle Z_t, w_t - w^\star\rangle \,|\, \mathcal{F}_{t-1}] = 0$$

and

$$\mathbb{E}\big[\, \|\widetilde{g}_t\|_2^2 \,\big|\, \mathcal{F}_{t-1}\big] = \|\nabla L_D(w_t)\|_2^2 + d\tau^2 \le G^2 + d\tau^2.$$

Then use these identities with part (a), taking $u = w^\star$, to show that

$$\langle \nabla L_D(w_t), w_t - w^\star\rangle \le \frac{\|w_t - w^\star\|_2^2 - \mathbb{E}[\|w_{t+1} - w^\star\|_2^2 \mid \mathcal{F}_{t-1}]}{2\eta} + \frac{\eta}{2}(G^2 + d\tau^2).$$

(c) **Expected excess loss [12 points]** Show that the expected excess loss is

$$\boxed{\mathbb{E}[L_D(\bar{w}_T)] - L_D(w^\star) \le \frac{R^2}{2\eta T} + \frac{\eta}{2}(G^2 + d\tau^2).}$$

Hint: Use convexity of L_D.

(d) **Privacy of the output [8 points]** Now use *replace-one adjacency*: $D, D' \in \mathcal{X}^n$ are neighboring if they differ in one record. Show that, uniformly over $w \in \mathcal{W}$,

$$\|\nabla L_D(w) - \nabla L_{D'}(w)\|_2 \le \Delta := \frac{2G}{n}.$$

For the privacy proof, analyze the stronger mechanism that releases the entire noisy-gradient transcript $(\widetilde{g}_1, \dots, \widetilde{g}_T)$. Using the given facts, prove that the returned $\bar{w}_T$ is

$$\boxed{\rho_T = \frac{T\Delta^2}{2\tau^2} = \frac{2TG^2}{n^2\tau^2}\text{-zCDP}.}$$

Problem 3. Gaussian RDP, zCDP, and a separation [30 points]

For probability measures P, Q and $\alpha > 1$, define the Rényi divergence

$$D_\alpha(P\|Q) = \frac{1}{\alpha - 1} \log\left(\int \left(\frac{\mathrm{d}P}{\mathrm{d}Q} \right)^\alpha \mathrm{d}Q \right)$$

when $P \ll Q$, and set $D_\alpha(P\|Q) = +\infty$ otherwise. A randomized mechanism M is $(\alpha, \varepsilon_\alpha)$-RDP if

$$D_\alpha(M(D)\|M(D')) \le \varepsilon_\alpha$$

for every ordered pair of neighboring datasets D, D'.

For the next definition, let $P = M(D)$, $Q = M(D')$, and let $Y \sim P$. When $P \ll Q$, define the privacy loss $L_{D,D'}(Y) = \log(\mathrm{d}P/\mathrm{d}Q)(Y)$; otherwise, interpret the relevant expectation below as $+\infty$. We say that M is ρ-zCDP if, for every ordered pair of neighboring datasets and every $\lambda > 0$,

$$\log \mathbb{E}\left[e^{\lambda L_{D,D'}(Y)} \right] \le \rho\lambda(\lambda + 1).$$

(a) **The Gaussian mechanism [15 points]** Let

$$P = \mathcal{N}(\mu, \sigma^2 I_d), \qquad Q = \mathcal{N}(\mu', \sigma^2 I_d).$$

For every $\alpha > 1$, derive

$$\boxed{D_\alpha(P\|Q) = \frac{\alpha\|\mu - \mu'\|_2^2}{2\sigma^2}.}$$

Hint: With $X = \mu' + \sigma Z$, $Z \sim \mathcal{N}(0, I_d)$, evaluate $\mathbb{E}_{X\sim Q}[(p(X)/q(X))^\alpha]$.

Now let f have ℓ_2-sensitivity at most Δ, and consider

$$M(D) = f(D) + \mathcal{N}(0, \sigma^2 I_d).$$

Conclude that, for every $\alpha > 1$,

$$M \text{ is } \left(\alpha, \frac{\alpha\Delta^2}{2\sigma^2} \right)\text{-RDP}.$$

(b) **RDP and zCDP [5 points]** Prove that:

$$\boxed{M \text{ is } \rho\text{-zCDP} \iff M \text{ is } (\alpha, \rho\alpha)\text{-RDP for every } \alpha > 1.}$$

Why is an RDP guarantee at only one order α not, by itself, a zCDP guarantee? Identify the zCDP parameter of the Gaussian mechanism in part (a).

(c) **Approximate DP need not imply finite zCDP [10 points]** Fix $\delta \in (0, 1)$. Prove that there is a mechanism that satisfies (ε, δ)-DP for every $\varepsilon \ge 0$, but does not satisfy ρ-zCDP for any finite value of ρ. *Hint:* Try proving the claim with a $(0, \delta)$-DP mechanism.