Week 9: Online Regret and Mirror Descent

Regularization leads to low regret in the online setting

Tianhao Wang

`tianhaowang@ucsd.edu`

UCSD · Spring 2026

DSC 190/291 Topics: Learning Theory

Contents

Bridge from Week 8

Same object, new interpretation

Week 8 RERM (batch, run **once** on a sample S of size n):

$$\hat{w}_\lambda = \arg\min_{w \in \mathcal{W}} L_S(w) + \lambda \Psi(w).$$

Week 9 FTRL (online, re-run at **every** round t on the data seen so far):

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \frac{1}{t} \sum_{i=1}^{t} \ell(w, z_i) + \lambda_t \Psi(w).$$

Same **rule** (regularized empirical loss minimization), re-invoked each round.

The new question is the **analysis**:

- Week 8 bounded $L_{\mathcal{D}}(\hat{w}_\lambda)$ on an i.i.d. sample.
- Week 9 bounds the **regret** $\frac{1}{n} \sum_{t=1}^{n} \ell(w_t, z_t) - \inf_{w \in \mathcal{W}} \frac{1}{n} \sum_{t=1}^{n} \ell(w, z_t)$ on an arbitrary sequence.

Regularization leads to low regret in the online setting.

Roadmap for Week 9

Part	Question	Main idea
1	What does online regret measure?	compete with best fixed action in hindsight
2	Why can FTL fail?	stability is not automatic
3	How does regularization fix FTL?	FTRL: strong convexity $\to$ stability $\to$ regret
4	How do we get cheap updates?	linearize: FTRLL, OGD
5	Beyond Euclidean geometry?	mirror descent; Ψ chooses the geometry
6	Finite hypothesis classes?	exponentiated gradient $\to$ multiplicative weights

Online Regret

From mistake bounds to regret

Week 1 (online binary classification): at each round $t = 1, 2, \ldots,$

- receive input $x_t \in \mathcal{X}$ (e.g. an email);
- predict $\hat{y}_t \in \{\pm 1\}$ using past data $(x_1, y_1), \ldots, (x_{t-1}, y_{t-1})$;
- observe the true label $y_t \in \{\pm 1\}$;
- count a mistake if $\hat{y}_t \neq y_t$.

Week 9 (general online learning): at each round t,

- choose a hypothesis $h_t \in \mathcal{H}$ using past examples $z_1, \ldots, z_{t-1}$;
- observe an example $z_t \in \mathcal{Z}$ (generalizes Week 1's pair (x_t, y_t));
- suffer loss $\ell(h_t, z_t) \in \mathbb{R}$, where $\ell : \mathcal{H} \times \mathcal{Z} \to \mathbb{R}$.

Two concrete settings for $(\mathcal{H}, \mathcal{Z}, \ell)$:

- **Supervised learning**: $z = (x, y)$, $h : \mathcal{X} \to \mathcal{Y}$ is a predictor, $\ell(h, (x, y)) = \mathrm{loss}(h(x), y)$.
- **Online investment**: $h \in \Delta_d$ is a portfolio over d stocks, $z \in \mathbb{R}^d$ is the day's market returns, $\ell(h, z) = -\langle h, z \rangle$ is the negative return.

Online regret: the definition

In Week 8 we judged a rule by its true risk $L_{\mathcal{D}}(h) = \mathbb{E}_z \ell(h, z)$.

Online sequences need not be i.i.d., so there is no $\mathcal{D}$ and no $L_{\mathcal{D}}$.

Instead, compare the learner's cumulative loss to the **best fixed action in hindsight**.

A learning rule produces $h_t = A(z_1, \ldots, z_{t-1}) \in \mathcal{H}$. Its **average regret** is

$$\mathrm{Reg}(n) = \frac{1}{n}\sum_{t=1}^{n} \ell(h_t, z_t) - \inf_{h \in \mathcal{H}} \frac{1}{n}\sum_{t=1}^{n} \ell(h, z_t).$$

- z_t may be **arbitrary**: no i.i.d. assumption, possibly adversarial.
- The comparator is **fixed across rounds** but chosen knowing the whole sequence.

Follow the Leader

The simplest online rule: play the h that would have been best on the past.

Follow the Leader (FTL), the online analog of ERM:

$$h_t = \arg\min_{h \in \mathcal{H}} \sum_{i=1}^{t-1} \ell(h, z_i).$$

To analyze FTL, compare it to a fictional rule that also sees z_t.

Be the Leader (BTL):

$$h_t^+ = \arg\min_{h \in \mathcal{H}} \sum_{i=1}^{t} \ell(h, z_i).$$

BTL has nonpositive regret

$$\sum_{t=1}^{n} \ell(h_t^+, z_t) \leq \inf_{h \in \mathcal{H}} \sum_{t=1}^{n} \ell(h, z_t).$$

Why BTL has nonpositive regret

Prove the BTL lemma by induction on n. Base $n = 0$: both sides equal 0.

Induction step: assume the bound for $n-1$ and apply with comparator h_n^+,

$$\begin{aligned}
\sum_{t=1}^{n-1} \ell(h_t^+, z_t) + \ell(h_n^+, z_n) &\le \sum_{t=1}^{n-1} \ell(h_n^+, z_t) + \ell(h_n^+, z_n) && \text{(induction with } h \leftarrow h_n^+\text{)} \\
&= \sum_{t=1}^{n} \ell(h_n^+, z_t) && \text{(combine the two terms)} \\
&\le \sum_{t=1}^{n} \ell(h, z_t) && \text{(optimality of } h_n^+\text{)}.
\end{aligned}$$

BTL beats every fixed h, but it is not implementable: it uses z_t before round t.

FTL chooses h_t from $z_1, \ldots, z_{t-1}$; BTL chooses h_t^+ from $z_1, \ldots, z_t$.

So FTL tracks BTL when adding z_t barely changes A's loss at z_t.

Leave-last-out stability (online variant of Week 8's leave-one-out):

$$|\ell(A(z_1, \ldots, z_t), z_t) - \ell(A(z_1, \ldots, z_{t-1}), z_t)| \leq \beta(t).$$

FTL via leave-last-out stability

If FTL is $\beta(t)$-stable, then

$$\mathrm{Reg}_{\mathsf{FTL}}(n) \leq \frac{1}{n} \sum_{t=1}^{n} \beta(t).$$

Proof of stability-to-regret

At round t, FTL plays $h_t = \arg\min_{h \in \mathcal{H}} \sum_{i=1}^{t-1} \ell(h, z_i)$.

FTL's next play h_{t+1} is also BTL's current play h_t^+:

$$h_{t+1} = \arg\min_{h \in \mathcal{H}} \sum_{i=1}^{t} \ell(h, z_i) = h_t^+.$$

Leave-last-out stability of FTL compares h_t and h_{t+1}:

$$\ell(h_t, z_t) \le \ell(h_{t+1}, z_t) + \beta(t) = \ell(h_t^+, z_t) + \beta(t).$$

Sum over t and apply the BTL lemma:

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^{n} \ell(h_t, z_t) &\le \frac{1}{n}\sum_{t=1}^{n} \ell(h_t^+, z_t) + \frac{1}{n}\sum_{t=1}^{n} \beta(t) && \text{(stability)} \\ &\le \inf_{h \in \mathcal{H}} \frac{1}{n}\sum_{t=1}^{n} \ell(h, z_t) + \frac{1}{n}\sum_{t=1}^{n} \beta(t) && \text{(BTL lemma).} \end{aligned}$$

Stable FTL is a low-regret online learner.

When FTL behaves well

When does FTL achieve stability? A natural case: squared-loss prediction.

$$\mathcal{H} = \mathbb{R}^d, \quad \mathcal{Z} = \{z : \|z\|_2 \leq 1\}, \quad \ell(h, z) = \|h - z\|_2^2.$$

Each round, the adversary picks a target z_t in the unit ball; the learner predicts h_t.

FTL predicts the empirical mean:

$$h_{t+1} = \frac{1}{t} \sum_{i=1}^{t} z_i.$$

We will show FTL is $\beta(t) = \frac{8}{t}$-stable, hence

$$\mathrm{Reg}_{\mathsf{FTL}}(n) \leq \frac{1}{n} \sum_{t=1}^{n} \frac{8}{t} = O\left(\frac{\log n}{n}\right).$$

> Curvature in the loss can make FTL stable even without an explicit regularizer.

Squared loss: the stability calculation

For $t \geq 2$, $h_t = \frac{1}{t-1} \sum_{i=1}^{t-1} z_i$ is the FTL output before round t.

h_{t+1} is the mean of $z_1, \ldots, z_t$, hence a convex combination of h_t and z_t:

$$h_{t+1} = \left(1 - \frac{1}{t}\right) h_t + \frac{1}{t} z_t, \quad h_{t+1} - z_t = \left(1 - \frac{1}{t}\right)(h_t - z_t).$$

Both $\|z_t\|_2 \leq 1$ and $\|h_t\|_2 \leq 1$ (a mean of unit-ball vectors), so $\|h_t - z_t\|_2 \leq 2$. Then

$$\begin{aligned} \left|\ell(h_{t+1}, z_t) - \ell(h_t, z_t)\right| &= \left| \|h_{t+1} - z_t\|_2^2 - \|h_t - z_t\|_2^2 \right| \\ &= \left(1 - \left(1 - \frac{1}{t}\right)^2\right) \cdot \|h_t - z_t\|_2^2 \\ &\leq \left(\frac{2}{t}\right) \cdot 4 = \frac{8}{t}. \end{aligned}$$

When FTL fails

Linear problem on the interval:

$$\mathcal{H} = [-1, 1], \quad \mathcal{Z} = [-1, 1], \quad \ell(h, z) = hz.$$

FTL plays $h_t = -\operatorname{sign}\left(\sum_{i<t} z_i\right)$, the opposite sign of the past total.

The adversary picks $z_1 = \frac{1}{2}$, $z_2 = -1$, $z_3 = 1$, $z_4 = -1$, ... so the past sum **flips sign** after each round.

- Learner: h_t and z_t have opposite signs, so $h_t z_t = 1$ for $t \geq 2$, giving $\sum_t h_t z_t = n - 1$.
- Comparator: partial sums satisfy $\left|\sum_{t=1}^{n} z_t\right| = \frac{1}{2}$, so $\inf_{h \in [-1,1]} h \sum_t z_t = -\frac{1}{2}$.

Hence $\operatorname{Reg}(n) = \frac{n - \frac{1}{2}}{n} \approx 1$: regret stays **constant per round**.

Convexity and Lipschitzness alone do **not** make FTL stable. We need explicit regularization.

Follow the Regularized Leader

FTRL: add a regularizer for stability

Recall the online setting:

- At round t, the learner picks $w_t \in \mathcal{W}$, observes $z_t \in \mathcal{Z}$, pays $\ell(w_t, z_t)$.
- $\mathcal{W} \subset \mathbb{R}^d$ is convex; $\ell(\cdot, z)$ is convex in w.

Follow the Regularized Leader adds a **strongly convex regularizer** Ψ to FTL:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \frac{1}{t} \sum_{i=1}^{t} \ell(w, z_i) + \lambda_t \Psi(w).$$

$\lambda_t > 0$ is a regularization schedule chosen below.

The mechanism mirrors Week 8 RERM:

RERM (batch, Week 8)

Ψ + i.i.d. $\rightarrow$ stability $\rightarrow$ generalization.

FTRL (online, Week 9)

Ψ + online $\rightarrow$ stability $\rightarrow$ low regret.

Recall the regret:

$$\mathrm{Reg}(n) = \frac{1}{n}\sum_{t=1}^{n} \ell(w_t, z_t) - \inf_{w\in\mathcal{W}} \frac{1}{n}\sum_{t=1}^{n} \ell(w, z_t).$$

FTRL regret bound

Suppose ℓ and Ψ are taken with respect to a common norm $\|\cdot\|$ on $\mathbb{R}^d$:

- $\ell(w, z)$ is convex and G-Lipschitz in w;
- Ψ is nonnegative and α-strongly convex, with $\sup_{w\in\mathcal{W}} \Psi(w) \le B^2$.

With the schedule $\lambda_t = \sqrt{\frac{G^2}{\alpha t B^2}}$, FTRL achieves

$$\mathrm{Reg}(n) = O\left(\frac{BG}{\sqrt{\alpha n}}\right).$$

Same regularizer Ψ and same $1/\sqrt{n}$ rate as Week 8, now on an arbitrary online sequence.

FTRL = FTL on a modified loss

Recall FTRL:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \frac{1}{t} \sum_{i=1}^{t} \ell(w, z_i) + \lambda_t \Psi(w).$$

We prove the regret bound first for **constant** λ, then extend.

With $\lambda_t = \lambda$, the FTRL objective equals $\frac{1}{t} \sum_{i=1}^{t} \tilde{\ell}(w, z_i)$ for the **modified loss**

$$\tilde{\ell}(w, z) = \ell(w, z) + \lambda \Psi(w).$$

So FTRL is FTL on $\tilde{\ell}$, which is $\lambda\alpha$-strongly convex (ℓ convex, Ψ is α-strongly convex).

Apply Week 8′s RERM stability bound at round t:

$$\beta(t) \leq \frac{2G^2}{\lambda \alpha t}.$$

FTRL inherits its per-round stability from Week 8′s RERM bound.

From stability to regret bound (constant λ)

Invoke the stability → regret lemma (FTL on $\tilde{\ell}$ = FTRL) with $\beta(t) = \frac{2G^2}{\lambda\alpha t}$:

$$\begin{aligned}
\mathrm{Reg}(n) &= \frac{1}{n}\sum_{t=1}^{n} \ell(w_t, z_t) - \inf_{w\in\mathcal{W}} \frac{1}{n}\sum_{t=1}^{n} \ell(w, z_t) && \text{(definition)}\\
&\leq \frac{1}{n}\sum_{t=1}^{n} \tilde{\ell}(w_t, z_t) - \inf_{w\in\mathcal{W}} \frac{1}{n}\sum_{t=1}^{n} \ell(w, z_t) && (\text{drop } \lambda\Psi(w_t) \geq 0)\\
&\leq \frac{1}{n}\sum_{t=1}^{n} \tilde{\ell}(w_t, z_t) - \inf_{w\in\mathcal{W}} \frac{1}{n}\sum_{t=1}^{n} \tilde{\ell}(w, z_t) + \lambda B^2 && (\text{bound } \lambda\Psi(w) \leq \lambda B^2)\\
&\leq \frac{1}{n}\sum_{t=1}^{n} \frac{2G^2}{\lambda\alpha t} + \lambda B^2 && \text{(stability lemma)}\\
&= \frac{2G^2 H_n}{\lambda\alpha n} + \lambda B^2 && \left(H_n = \sum_t \frac{1}{t} = O(\log n)\right).
\end{aligned}$$

Constant λ gives $\mathrm{Reg}(n) \leq \lambda B^2 + O\left(\frac{G^2 \log n}{\lambda\alpha n}\right)$.

Balancing the FTRL terms

Choose λ to minimize the bound $\lambda B^2 + \frac{G^2 \log n}{\lambda \alpha n}$. The two terms balance when

$$\lambda B^2 = \frac{G^2 \log n}{\lambda \alpha n} \quad \Rightarrow \quad \lambda = \left(\frac{G}{B}\right)\sqrt{\frac{\log n}{\alpha n}}.$$

Plugging back,

$$\mathrm{Reg}(n) \leq O\left(BG\sqrt{\frac{\log n}{\alpha n}}\right).$$

Compared to the target $O(BG/\sqrt{\alpha n})$, this has an extra $\sqrt{\log n}$ factor. A time-varying λ_t removes it.

Refined FTRL: time-varying λ_t

With λ_t varying in t, FTRL at round t minimizes

$$\frac{1}{t}\sum_{i=1}^{t}\ell(w,z_i)+\lambda_t\Psi(w).$$

To express this as FTL on round-dependent losses $\tilde{\ell}_i(w,z)=\ell(w,z)+\lambda_i'\Psi(w)$, we need

$$\frac{1}{t}\sum_{i=1}^{t}\tilde{\ell}_i(w,z_i)=\frac{1}{t}\sum_{i=1}^{t}\ell(w,z_i)+\left(\frac{1}{t}\sum_{i=1}^{t}\lambda_i'\right)\Psi(w)$$

to equal the FTRL objective, i.e., $\sum_{i=1}^{t}\lambda_i'=t\lambda_t$. Telescoping gives $\lambda_t'=t\lambda_t-(t-1)\lambda_{t-1}$.

Leave-last-out stability gives $\beta(t)\le\frac{2G^2}{\lambda_t\alpha t}$. With $\lambda_t=\sqrt{\frac{G^2}{\alpha t B^2}}$, so $\beta(t)=\frac{2GB}{\sqrt{\alpha t}}$ and $\sum_{t=1}^{n}\frac{1}{\sqrt{t}}=O(\sqrt{n})$:

$$\mathrm{Reg}(n)\le\lambda_n B^2+\frac{1}{n}\sum_{t=1}^{n}\beta(t)\le\frac{GB}{\sqrt{\alpha n}}+\frac{2GB}{n\sqrt{\alpha}}\sum_{t=1}^{n}\frac{1}{\sqrt{t}}\le O\left(\frac{BG}{\sqrt{\alpha n}}\right).$$

The right schedule removes the $\sqrt{\log n}$ factor and matches the Week 8 rate.

FTRL is expensive per round

FTRL achieves the right regret rate but is computationally awkward.

Per-round cost. At round t, FTRL solves $\arg\min_{w \in \mathcal{W}} \frac{1}{t} \sum_{i=1}^{t} \ell(w, z_i) + \lambda_t \Psi(w)$.

- Work per round: scales with t (a sweep over all $t - 1$ examples).
- Cumulative work over n rounds: $\Omega(n^2)$.

Memory. The objective requires the full history at every round.

- Stores $z_1, \ldots, z_{t-1}$.
- State size at round t: $\Theta(t)$.

We would like an online method that does $O(1)$ **work and** $O(1)$ **extra memory per round**, not $O(t)$.

Linearization: From FTRL to OGD

Linear losses are easy

For **linear losses** $\ell(w, g) = \langle g, w \rangle$ where $\|g\|_* \le G$, FTRL becomes

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \frac{1}{t} \left\langle \sum_{i=1}^{t} g_i, w \right\rangle + \lambda_t \Psi(w).$$

Take $\Psi(w) = \frac{1}{2}\|w\|_2^2$ and $\mathcal{W} = \mathbb{R}^d$ (no constraints). First-order condition:

$$0 = \frac{1}{t} \sum_{i=1}^{t} g_i + \lambda_t w_{t+1} \quad \Rightarrow \quad w_{t+1} = -\frac{1}{t\lambda_t} \sum_{i=1}^{t} g_i.$$

Rewrite recursively (multiply and divide by previous coefficient):

$$w_{t+1} = \frac{\lambda_{t-1}(t-1)}{\lambda_t t} w_t - \frac{1}{\lambda_t t} g_t.$$

Linear losses give a cheap recursive update: just maintain the cumulative gradient.

Two pieces in hand:

- Linear FTRL: cheap $O(1)$-per-round update, regret $O\left(B\frac{G}{\sqrt{\alpha n}}\right)$.
- General convex FTRL: $O(n^2)$ work, $\Theta(t)$ memory per round.

Plan. Reduce general convex losses to the linear case:

1. bound each convex $\ell(\cdot, z_t)$ below by a linear function of w;
2. show regret on ℓ is bounded by regret on those linear surrogates;
3. apply linear FTRL to the surrogates.

The engine. Assume $\ell(\cdot, z)$ is **differentiable** and convex. Then for every w,

$$\ell(w, z_t) \geq \ell(w_t, z_t) + \langle g_t, w - w_t \rangle, \quad g_t = \nabla_w \ell(w_t, z_t).$$

Convexity hands us a free linear lower bound at w_t.

Linearizing convex losses

Set $g_t = \nabla_w \ell(w_t, z_t)$ and define the **linearized loss** at round t as

$$\tilde{\ell}_t(w) = \langle g_t, w \rangle.$$

If ℓ is G-Lipschitz w.r.t. $\|\cdot\|$, then $\|g_t\|_* \leq G$.

Linearization reduction

If $\ell(\cdot, z)$ is convex and differentiable in w for every z, then for every $w \in \mathcal{W}$,

$$\sum_{t=1}^{n} \ell(w_t, z_t) - \sum_{t=1}^{n} \ell(w, z_t) \leq \sum_{t=1}^{n} \langle g_t, w_t \rangle - \sum_{t=1}^{n} \langle g_t, w \rangle.$$

Proof. By the tangent-plane lower bound $\ell(w, z_t) \geq \ell(w_t, z_t) + \langle g_t, w - w_t \rangle$,

$$\ell(w_t, z_t) - \ell(w, z_t) \leq \langle g_t, w_t - w \rangle = \langle g_t, w_t \rangle - \langle g_t, w \rangle.$$

Sum over $t = 1, \ldots, n$ to obtain the claim.

Setting. $\ell(\cdot, z)$ is **convex**, differentiable, and G-Lipschitz w.r.t. $\|\cdot\|$ for every z.

Definition. Apply FTRL to the linearized losses $\tilde{\ell}_t(w) = \langle g_t, w \rangle$:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \frac{1}{t} \left\langle \sum_{i=1}^{t} g_i, w \right\rangle + \lambda_t \Psi(w).$$

Call this rule **FTRLL** (Follow the Regularized Linearized Leader).

Regret bound. Chain two facts:

$$\begin{aligned} \mathrm{Reg}_{\mathsf{FTRLL}\text{ on }\ell}(n) &\le \mathrm{Reg}_{\mathsf{FTRL}\text{ on }\tilde{\ell}}(n) && \text{(linearization reduction)} \\ &\le O\left(\frac{BG}{\sqrt{\alpha n}}\right) && \text{(FTRL bound; } \tilde{\ell}_t \text{ is } G\text{-Lipschitz since } \|g_t\|_* \le G). \end{aligned}$$

Computational benefit.

- History compressed into the **cumulative gradient** $\sum_{i \le t} g_i$.
- State per round: $O(d)$, not $O(t)$.

Recall FTRLL. For Euclidean potential $\Psi(w) = \frac{1}{2}\|w\|_2^2$ on $\mathcal{W} = \mathbb{R}^d$,

$$w_{t+1} = \frac{\lambda_{t-1}(t-1)}{\lambda_t t} w_t - \frac{1}{\lambda_t t} g_t.$$

Two natural schedules ($g_t = \nabla \ell(w_t, z_t)$, $\|g_t\|_2 \le G$, $\|w\|_2 \le B$):

Time-varying, $\lambda_t = \lambda/\sqrt{t}$.

- Update: $w_{t+1} = \sqrt{1 - \frac{1}{t}} w_t - g_t / \left(\lambda\sqrt{t}\right)$.
- Regret: $\mathrm{Reg}(n) \le O\left(\frac{BG}{\sqrt{n}}\right)$, optimal rate, anytime.

Constant stepsize, $\lambda_t = \lambda/t$.

- Then $t\lambda_t = \lambda$, so first coefficient $= 1$ and second $= \frac{1}{\lambda}$ exactly.
- Update: $w_{t+1} = w_t - g_t/\lambda$. This is **online gradient descent (OGD)** with constant stepsize $\eta = \frac{1}{\lambda}$.
- Regret: $\mathrm{Reg}(n) \le O\left(\frac{BG}{\sqrt{n}}\right)$ when λ is tuned to horizon n.

Perceptron as OGD

Recall the Perceptron update from Week 1:

$$w_{t+1} = w_t + \mathbf{1}[y_t \langle w_t, x_t \rangle \leq 0] \cdot y_t x_t.$$

This is OGD with stepsize 1 on the **zero-margin hinge**

$$\ell(w, (x, y)) = \max(0, -y\langle w, x \rangle).$$

The **aggressive Perceptron** updates whenever the margin is below 1:

$$w_{t+1} = w_t + \mathbf{1}[y_t \langle w_t, x_t \rangle \leq 1] \cdot y_t x_t,$$

which is OGD on the standard hinge $\max(0, 1 - y\langle w, x \rangle)$.

The hinge is non-differentiable at the kink $y\langle w, x \rangle = 1$; we use the left-derivative there, which equals $-yx$ when the hinge is active and 0 otherwise. Treating the kink with the left-derivative makes the Perceptron update exactly the gradient step.

> Week 1's Perceptron is the Euclidean special case of today's framework.

FTRLL regret bound

FTRLL for convex Lipschitz bounded problems

Assumptions.

- **Setup.** $\mathcal{B} = \mathbb{R}^d$ ambient normed space; $\mathcal{W} \subseteq \mathcal{B}$ convex comparator set.
- **Regularizer.** $\Psi : \mathcal{B} \to \mathbb{R}_{\geq 0}$ is α-strongly convex w.r.t. $\|\cdot\|$, with $\sup_{w \in \mathcal{W}} \Psi(w) \leq B^2$.
- **Loss.** For every $z \in \mathcal{Z}$, $\ell(\cdot, z) : \mathcal{B} \to \mathbb{R}$ is convex, differentiable, and G-Lipschitz w.r.t. $\|\cdot\|$ (equivalently, $\|\nabla_w \ell(w, z)\|_* \leq G$).

Conclusion. FTRLL with schedule $\lambda_t = \sqrt{\frac{G^2}{\alpha t B^2}}$ achieves

$$\mathrm{Reg}(n) \leq O\left(\frac{BG}{\sqrt{\alpha n}}\right).$$

- Same regret rate $O\left(\frac{BG}{\sqrt{\alpha n}}\right)$ as FTRL.
- FTRL stores all t past examples; FTRLL compresses them into the **cumulative gradient** $\sum_{i \leq t} g_i$.

Linearization buys computational efficiency at no statistical cost.

Choosing the regularizer

FTRLL across choices of Ψ

FTRLL works for **any** α-strongly convex Ψ, with regret

$$\mathrm{Reg}(n) \le O\left(\frac{BG}{\sqrt{\alpha n}}\right).$$

All three constants in the bound depend on Ψ through a common norm $\|\cdot\|$:

- B, comparator radius: $\sup_{w \in \mathcal{W}} \Psi(w) \le B^2$;
- α, strong-convexity constant of Ψ w.r.t. $\|\cdot\|$;
- G, Lipschitz constant of ℓ w.r.t. $\|\cdot\|$.

We've used $\Psi(w) = \frac{1}{2}\|w\|_2^2$ (OGD). Three other choices follow:

1. **preconditioned OGD** (quadratic Ψ),
2. ℓ_p **geometry**,
3. **exponentiated gradient** on the simplex (entropic Ψ).

Preconditioned OGD

FTRLL with $\Psi(w) = \frac{1}{2} w^\top Q w$, Q positive definite.

Ψ is 1-strongly convex w.r.t. $\|w\|_Q = \sqrt{w^\top Q w}$; dual norm $\|v\|_{Q,*} = \sqrt{v^\top Q^{-1} v}$.

FTRLL closed form. The first-order condition for $\arg\min \frac{1}{t} \left\langle \sum_{i=1}^t \nabla \ell_i, w \right\rangle + \lambda_t \Psi(w)$ gives

$$w_{t+1} = -\frac{1}{t\lambda_t} Q^{-1} \sum_{i=1}^t \nabla \ell_i.$$

With the constant-stepsize schedule $\lambda_t = \lambda/t$ (as in the OGD slide), this is recursively

$$w_{t+1} = w_t - \eta Q^{-1} \nabla \ell_t, \quad \eta = 1/\lambda.$$

This is **preconditioned OGD**: Q^{-1} rescales the gradient step.

Regret (with $\|w\|_Q \leq B$):

$$\mathrm{Reg}(n) \leq O\left(\frac{B \|\nabla \ell\|_{Q,*}}{\sqrt{n}}\right).$$

ℓ_p geometry

FTRLL with $\Psi(w) = \frac{1}{2}\|w\|_p^2$ for $1 < p \le 2$.

Ψ is **$(p-1)$-strongly convex** w.r.t. $\|\cdot\|_p$, with gradient $\nabla\Psi(w)_i = \|w\|_p^{2-p}|w_i|^{p-1}\,\mathrm{sign}(w_i)$.

FTRLL ($\lambda_t = \lambda/t$), first-order condition:

$$\nabla\Psi(w_{t+1}) = -\frac{1}{\lambda}\sum_{i\le t}\nabla\ell_i.$$

Regret (with $\|w\|_p \le B$):

$$\mathrm{Reg}(n) \le O\left(\frac{B\|\nabla\ell\|_q}{\sqrt{(p-1)n}}\right).$$

Bound **explodes as $p \to 1$**. For **ℓ_1-bounded** comparators (e.g. the simplex), two options:

- tune $p = 1 + O(1/\log d)$, giving $\|\nabla\ell\|_\infty\sqrt{\log d/n}$;
- use a Ψ suited to $\|\cdot\|_1$: the entropic potential (next).

FTRLL with the entropic potential on the simplex $\mathcal{W} = \Delta_d = \left\{ w \in \mathbb{R}^d : w \geq 0, \sum_i w_i = 1 \right\}$:

$$\Psi(w) = \sum_{i=1}^{d} w_i \log(dw_i) = \log d + \sum_{i=1}^{d} w_i \log w_i .$$

Plug into the FTRLL theorem (work in $\|\cdot\|_1$, dual $\|\cdot\|_\infty$):

- $0 \leq \Psi(w) \leq \log d$, so $B^2 = \log d$;
- Ψ is **1-strongly convex w.r.t.** $\|\cdot\|_1$ (Pinsker, as in Week 8), so $\alpha = 1$;
- gradients measured in dual norm $\|\cdot\|_\infty$.

Conclusion: $\mathrm{Reg}(n) \leq O\left(\sqrt{\frac{\|\nabla \ell\|_\infty \log d}{n}} \right)$.

From $\nabla\Psi(w)_i = \log w_i + 1 + \log d$ we get $\nabla\Psi^{-1}(\nu)_i = e^{\nu_i - 1 - \log d}$. The FTRLL minimizer on Δ_d:

$$w_{t+1}[i] \propto \exp\left(- \frac{\left(\sum_{s \leq t} \nabla \ell_s \right)[i]}{t \lambda_t} \right).$$

Exponentiated gradient

Specialize $\lambda_t = \lambda/t$, $\eta = 1/\lambda$ (so $t\lambda_t = \lambda$). The entropic FTRLL update becomes

$$w_{t+1}[i] \propto \exp\left(-\eta\left(\sum_{s\le t}\nabla\ell_s\right)[i]\right) = \prod_{s\le t}\exp(-\eta(\nabla\ell_s)[i]).$$

Pull out the $s = t$ factor and renormalize on Δ_d to get the **multiplicative recursive update**:

$$w_{t+1}[i] = \frac{w_t[i]\cdot\exp(-\eta(\nabla\ell_t)[i])}{\sum_j w_t[j]\cdot\exp(-\eta(\nabla\ell_t)[j])}, \quad w_1[i] = \frac{1}{d}.$$

This is the **Exponentiated Gradient (EG)** algorithm.

Comparison: OGD on the simplex. OGD ($\Psi = \frac{1}{2}\|\cdot\|_2^2$) measures in $\|\cdot\|_2$, not $\|\cdot\|_1$:

- $\|w\|_2 \le \|w\|_1 = 1$, so $B = 1$;
- norm conversion: $\|\nabla\ell\|_2 \le \sqrt{d}\|\nabla\ell\|_\infty$;
- FTRLL theorem $\Rightarrow$ $\mathrm{Reg}(n) \le O\left(\|\nabla\ell\|_\infty\sqrt{d/n}\right)$, vs $\|\nabla\ell\|_\infty\sqrt{(\log d)/n}$ for EG.

Matching the geometry of $\mathcal{W}$ replaces $\sqrt{d}$ by $\sqrt{\log d}$.

Multiplicative weights

Apply EG to a **finite class** $\mathcal{H}$, $|\mathcal{H}| = d$, loss in $[0, 1]$.

Play a distribution over $\mathcal{H}$. At round t: $w_t \in \Delta_d$, sample $h \sim w_t$, predict $h(x_t)$.

Expected loss is linear in w.

- Loss vector $\nabla \ell_t \in [0,1]^d$: $(\nabla \ell_t)[h] = \mathrm{loss}(h(x_t), y_t)$.
- Expected loss: $\mathbb{E}_{h \sim w_t} \mathrm{loss}(h(x_t), y_t) = \langle w_t, \nabla \ell_t \rangle$ (gradient $\nabla \ell_t$).

EG specializes to the **Multiplicative Weights** update:

$$w_{t+1}[h] \propto w_t[h] \cdot \exp(-\eta (\nabla \ell_t)[h]).$$

$\|\nabla \ell_t\|_\infty \leq 1$ (loss in $[0, 1]$) $\Rightarrow \mathrm{Reg}(n) \leq O\left(\sqrt{(\log d)/n}\right)$.

Week 1: Halving (realizable)

Total mistakes $\leq \log_2 d$.

Week 9: MW (agnostic)

Average regret $O\left(\sqrt{(\log d)/n}\right)$.

Both pay only $\log d$ in dimension; MW is the agnostic counterpart of Halving.

Recall the three FTRLL updates from our examples:

$$w_{t+1} = -\frac{1}{t\lambda_t}\sum_{i\le t}\nabla\ell_i \qquad \text{(Euclidean } \Psi\text{)}$$

$$w_{t+1} = -\frac{1}{t\lambda_t}Q^{-1}\sum_{i\le t}\nabla\ell_i \qquad \text{(quadratic } \Psi\text{)}$$

$$w_{t+1}[i] \propto \exp\left(-\frac{1}{t\lambda_t}\left(\sum_{s\le t}\nabla\ell_s\right)[i]\right) \qquad \text{(entropic } \Psi\text{)}$$

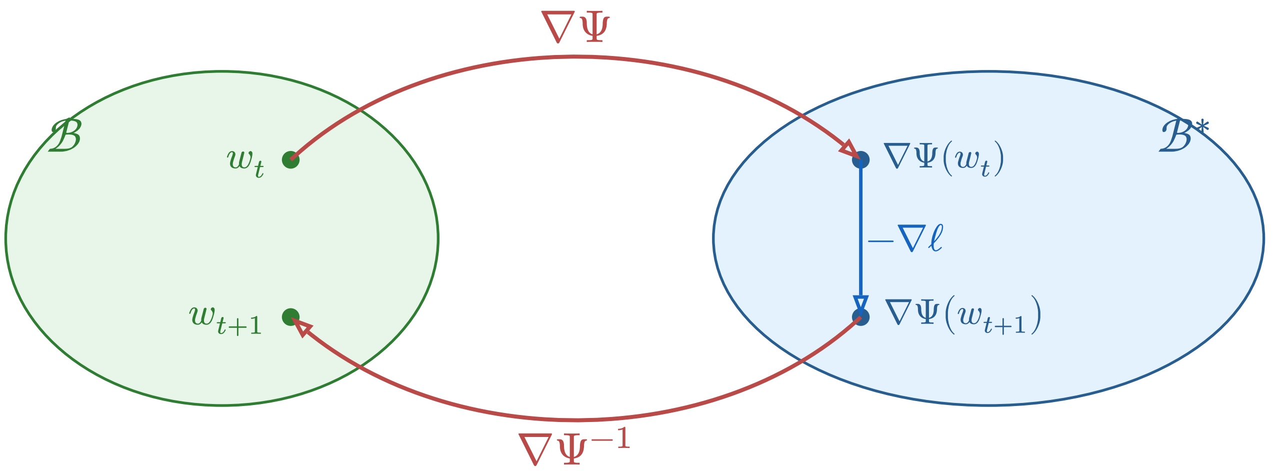
Derivation. FTRLL $\arg\min \frac{1}{t}\left\langle \sum_{i\le t}\nabla\ell_i, w\right\rangle + \lambda_t\Psi(w) \Rightarrow \nabla\Psi(w_{t+1}) = -\frac{1}{t\lambda_t}\sum_{i\le t}\nabla\ell_i$.

Ψ is ***α*-strongly convex** $\Rightarrow$ $\nabla\Psi$ is **strictly monotone hence injective** denote its inverse by $\nabla\Psi^{-1}$

$$w_{t+1} = \nabla\Psi^{-1}\left(-\frac{1}{t\lambda_t}\sum_{i\le t}\nabla\ell_i\right) = \nabla\Psi^{-1}\left(\frac{(t-1)\lambda_{t-1}}{t\lambda_t}\nabla\Psi(w_t) - \frac{1}{t\lambda_t}\nabla\ell_t\right)$$

The three cases: $\nabla\Psi^{-1}$ is identity, $Q^{-1}(\cdot)$, exponentiation.

Dual averaging: primal/dual picture



Each round traces a cycle through both spaces:

1. $w_t \to \nabla\Psi(w_t)$: **forward map $\nabla\Psi$** (primal $\to$ dual);
2. **accumulate the gradient** in dual: $\nabla\Psi(w_t) \to \nabla\Psi(w_{t+1})$;
3. $\nabla\Psi(w_{t+1}) \to w_{t+1}$: **inverse map $\nabla\Psi^{-1}$** (dual $\to$ primal).

Dual averaging: the dual coordinate $\nabla\Psi(w_{t+1})$ is the **running sum** $-\frac{1}{t\lambda_t}\sum_{i\le t}\nabla\ell_i$.

(Picture assumes $\mathcal{W} = \mathcal{B}$. For $\mathcal{W} \subsetneq \mathcal{B}$, project back to $\mathcal{W}$ via Bregman projection.)

Two ways to update the dual coordinate in the forward/dual picture:

FTRLL. Set it to the running sum (dual averaging):

$$\nabla\Psi(w_{t+1}) = -\frac{1}{t\lambda_t}\sum_{i\le t}\nabla\ell_i.$$

Mirror descent. Take an **incremental gradient step** in dual:

$$\nabla\Psi(w_{t+1}) = \nabla\Psi(w_t) - \eta_t\nabla\ell_t \quad\Leftrightarrow\quad \nabla\Psi(w_{t+1}) = -\sum_{i\le t}\eta_i\nabla\ell_i.$$

Both then apply $\nabla\Psi^{-1}$ to recover w_{t+1}.

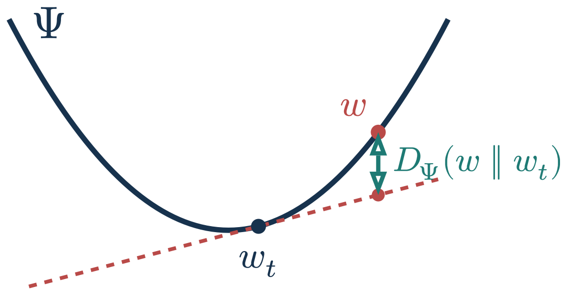
- Coincide for $\lambda_t = \lambda/t$ (constant $\eta_t = 1/\lambda$); otherwise distinct.
- **MD**: each $\nabla\ell_i$ carries its own weight η_i.
- **FTRLL**: all gradients share one weight $1/(t\lambda_t)$.

Bregman divergence

Given a differentiable convex potential Ψ, the **Bregman divergence** is

$$D_\Psi(w \parallel u) = \Psi(w) - \Psi(u) - \langle \nabla\Psi(u), w - u \rangle.$$

It measures how much Ψ at w exceeds its tangent plane at u.



- $D_\Psi(w \parallel u) \geq 0$ for convex Ψ, with equality iff $w = u$ (strict convexity).
- For $\Psi(w) = \frac{1}{2}\|w\|_2^2$, $D_\Psi(w \parallel u) = \frac{1}{2}\|w - u\|_2^2$.
- α-strong convexity of Ψ in $\|\cdot\|$ implies $D_\Psi(w \parallel u) \geq \frac{\alpha}{2}\|w - u\|^2$.

Proximal identity (motivates the MD update):

$$\arg\min_{x \in \mathcal{B}} \langle v, x \rangle + D_\Psi(x \parallel y) = \nabla\Psi^{-1}(\nabla\Psi(y) - v).$$

Online mirror descent

Non-linearized MD. Replace FTRL's history sum with a Bregman penalty toward w_t:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \ell(w, z_t) + \frac{1}{\eta_t} D_\Psi(w \,\|\, w_t).$$

Linearized MD. Replace $\ell(w, z_t)$ by its linear approximation $\langle \nabla \ell_t, w \rangle$:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \langle \nabla \ell_t, w \rangle + \frac{1}{\eta_t} D_\Psi(w \,\|\, w_t).$$

By the proximal identity, the unconstrained minimizer $\nabla \Psi^{-1}(\nabla \Psi(w_t) - \eta_t \nabla \ell_t) \in \mathcal{B}$ may be outside $\mathcal{W}$

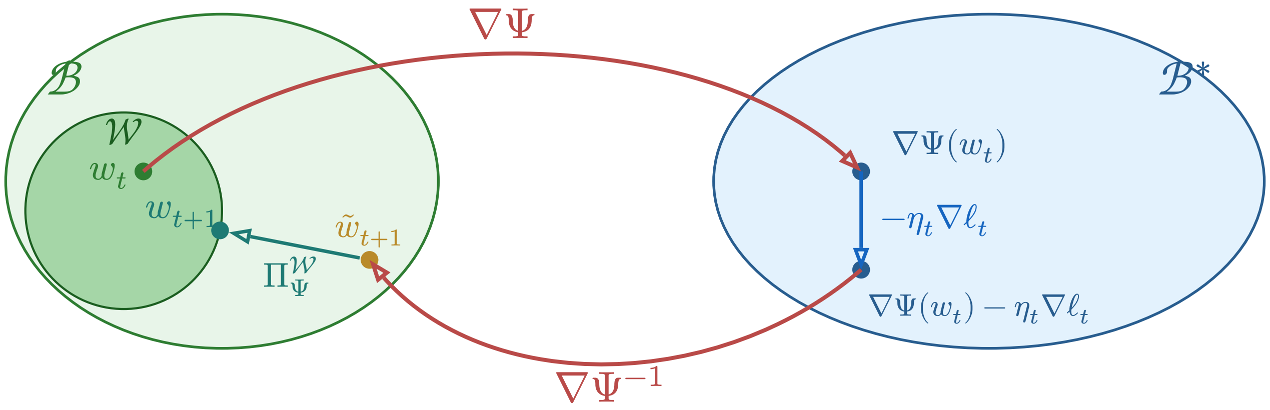
Project onto $\mathcal{W}$ using **Bregman projection** $\Pi_\Psi^{\mathcal{W}}(u) := \arg\min_{w \in \mathcal{W}} D_\Psi(w \,\|\, u)$:

$$w_{t+1} = \Pi_\Psi^{\mathcal{W}}(\nabla \Psi^{-1}(\nabla \Psi(w_t) - \eta_t \nabla \ell_t)).$$

MD achieves the **same regret rate** as FTRLL:

$$\mathrm{Reg}(n) \leq O\left(\sqrt{\frac{(\sup_{w \in \mathcal{W}} \Psi(w)) \cdot \|\nabla \ell\|_*^2}{\alpha n}}\right).$$

Mirror descent: primal $\leftrightarrow$ dual picture



Each MD round traces a cycle through both spaces, then projects back to $\mathcal{W}$:

1. $w_t \to \nabla\Psi(w_t)$: **forward map $\nabla\Psi$** (primal $\to$ dual);
2. **gradient step** in dual: $\nabla\Psi(w_t) \to \nabla\Psi(w_t) - \eta_t\nabla\ell_t$;
3. **inverse map $\nabla\Psi^{-1}$** back to $\mathcal{B}$: candidate $\tilde{w}_{t+1}$;
4. **Bregman projection** onto $\mathcal{W}$: $w_{t+1} = \Pi_\Psi^{\mathcal{W}}(\tilde{w}_{t+1})$.

Same forward/dual cycle as dual averaging, with an extra Bregman projection step.

Euclidean MD = projected gradient descent

Set $\Psi(w) = \frac{1}{2}\|w\|_2^2$. Then $D_\Psi(w \,\|\, w_t) = \frac{1}{2}\|w - w_t\|_2^2$ and $\nabla\Psi(w) = w$.

Non-linearized MD:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \ell(w, z_t) + \frac{1}{2\eta_t}\|w - w_t\|_2^2.$$

Linearized MD ≡ projected gradient descent:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \langle \nabla\ell_t, w\rangle + \frac{1}{2\eta_t}\|w - w_t\|_2^2 = \Pi_2^{\mathcal{W}}(w_t - \eta_t \nabla\ell_t),$$

where $\Pi_2^{\mathcal{W}}(w) := \arg\min_{w' \in \mathcal{W}} \|w - w'\|_2$ is the **Euclidean projection** onto $\mathcal{W}$.

Regret bound (specializing the MD bound with $\Psi = \frac{1}{2}\|\cdot\|_2^2$, $\alpha = 1$):

$$\mathrm{Reg}(n) \le O\left(\sqrt{\frac{(\sup_{w \in \mathcal{W}} \|w\|_2^2) \cdot \|\nabla\ell\|_2^2}{n}}\right).$$

Both fit the forward/dual picture; they differ in the dual update.

Linearized MD:

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \langle \nabla \ell_t, w \rangle + \frac{1}{\eta_t} D_\Psi(w \parallel w_t) = \Pi_\Psi^{\mathcal{W}}(\nabla \Psi^{-1}(\nabla \Psi(w_t) - \eta_t \nabla \ell_t)).$$

Contrast with linearized FTRL (dual averaging):

$$\begin{aligned} w_{t+1} &= \Pi_\Psi^{\mathcal{W}}\left(\nabla \Psi^{-1}\left(-\frac{1}{t\lambda_t}\sum_{i=1}^{t} \nabla \ell_i\right)\right) \\ &= \nabla \Psi^{-1}\left(\frac{(t-1)\lambda_{t-1}}{t\lambda_t} \nabla \Psi(w_t) - \frac{1}{t\lambda_t} \nabla \ell_t\right) \quad \text{(only if } \mathcal{W} = \mathcal{B}\text{)} \end{aligned}$$

Coincide with MD only when $\lambda_t = \lambda/t$ (then the coefficient on $\nabla\Psi(w_t)$ equals 1, and $\eta_t = 1/\lambda$).

- **Arbitrary stepsizes** without scaling of iterates: MD's update is recursive in w_t for any η_t.
- **Incremental algorithm** (keep only w_t, not the entire history), *even without linearization*: non-linearized MD uses the full loss $\ell(w, z_t) + \frac{1}{\eta_t} D_\Psi(w \parallel w_t)$.
- Can also **partially linearize**, or use any other convex lower bound on the objective.

Example: smooth loss + ℓ_1 penalty. At round t with example (x_t, y_t), let $f_t(w) = \mathrm{loss}(\langle w, x_t \rangle, y_t)$; the loss is $\ell(w, z_t) = f_t(w) + \|w\|_1$.

- Non-linearized: $w_{t+1} = \arg\min_{w \in \mathcal{W}} f_t(w) + \|w\|_1 + \frac{1}{\eta_t} D_\Psi(w \parallel w_t)$.
- Partially linearized (linearize f_t, keep ℓ_1):

$$w_{t+1} = \arg\min_{w \in \mathcal{W}} \langle \nabla f_t(w_t), w \rangle + \|w\|_1 + \frac{1}{\eta_t} D_\Psi(w \parallel w_t).$$

Useful for **sparse learning**, where the ℓ_1 structure should be preserved.

Online and Statistical Learning

Online and statistical

For convex G-Lipschitz problems with α-strongly convex Ψ satisfying $\sup_{w\in\mathcal{W}} \Psi(w) \le B^2$, the **same regret bound**

$$\mathrm{Reg}(n) \le O\left(\sqrt{\frac{B^2G^2}{\alpha n}}\right)$$

- FTRL, linearized FTRL (dual averaging);
- Mirror Descent, linearized MD (e.g., projected GD, EG), partially linearized MD.

Under the **same condition**, using stability arguments, RERM **agnostically PAC learns** with sample complexity $O\left(\frac{B^2G^2}{\alpha\varepsilon^2}\right)$ in the statistical setting.

Setting	Online rule	Batch / statistical analogue
finite class	Multiplicative Weights / Halving	ERM + union bound
ℓ_2 convex Lipschitz	OGD	RERM with $\lVert w\rVert_2^2$
simplex / ℓ_1 convex	Exponentiated Gradient	RERM with entropy
general Ψ	FTRL / Mirror Descent	RERM with Ψ

Summary

Week 9 in one picture

FTL (natural; can fail on linear losses)

$\downarrow$ add strongly convex Ψ

FTRL ($\Psi \Rightarrow$ stability $\Rightarrow$ Reg $= O(BG/\sqrt{\alpha n})$)

$\downarrow$ replace ℓ by $\langle g_t, \cdot \rangle$ (linearize)

FTRLL (cumulative-gradient update)

$\downarrow$ choose geometry via Ψ

OGD Mirror Descent EG / MW

Regularization leads to low regret in the online setting the choice of Ψ controls the geometry.

Final project

Weekly presentations are done. The final report bundles three artifacts, due **June 12** (end of quarter):

1. **Formalized proof.** Your Lean files, with a brief README pointing to the main theorem and noting what is assumed vs. proved.
2. **Designed proof presentation.** The creative core of the project. A good presentation of a proof lives somewhere between **natural-language exposition** (textbook style) and **machine-checked formalization** (your Lean code); your task is to design one that draws on both. Format is open: a webpage, an interactive notebook, a written essay, a tutorial document. Make the choices that serve your result and your reader.
3. **AI workflow.** The artifacts you built up across the whole quarter (homeworks + project): your `CLAUDE.md` / `AGENTS.md`, any subagents or skills. Add a short reflection on how the collaboration evolved over the term, what you delegated, what worked, and where AI misled you.

Submit by email to `tianhaowang@ucsd.edu`.