Adaptive online learning for games, optimization and deviation bounds

Karthik Sridharan
Cornell University

- based on work with Dylan Foster and Alexander Rakhlin
For $t = 1$ to $n$

- Adversary picks input instance $x_t \in \mathcal{X}$
- Learner picks prediction $\hat{y}_t \in \hat{\mathcal{Y}}$
- Adversary simultaneously picks label/output $y_t \in \mathcal{Y}$
- Learner suffers loss $\ell(\hat{y}_t, y_t)$

End
For $t = 1$ to $n$

Adversary picks input instance $x_t \in \mathcal{X}$
Learner picks prediction $\hat{y}_t \in \hat{\mathcal{Y}}$
Adversary simultaneously picks label/output $y_t \in \mathcal{Y}$
Learner suffers loss $\ell(\hat{y}_t, y_t)$

End

Goal: minimize regret

$$\text{Reg}_n = \sum_{t=1}^{n} \ell(\hat{y}_t, y_t) - \inf_{f \in \mathcal{F}} \sum_{t=1}^{n} \ell(f(x_t), y_t)$$
Adaptive Regret Bounds

- Typically we provide online learning algorithms
- Prove uniform bound on regret against worst case adversary
  \[ \text{Reg}_n \leq \text{Rate}(n) \]
- Can we get better bounds against nicer adversaries?
- And maintain the worst case guarantees?
Adaptive bound $B$:

\[ \forall f \in \mathcal{F}, x_{1:n}, y_{1:n}, \quad \sum_{t=1}^{n} \ell(\hat{y}_t, y_t) \leq \inf_{f \in \mathcal{F}} \left\{ \sum_{t=1}^{n} \ell(f(x_t), y_t) + B(x_{1:n}, y_{1:n}; f) \right\} \]

where $B$ is such that, $\sup_{f \in \mathcal{F}, x_{1:n}, y_{1:n}} B(x_{1:n}, y_{1:n}; f) \leq O\left(\text{Rate}(n)\right)$.
Adaptive Online Learning and Games
Both players honest: \( -\minimax \text{ equilibrium at rate } -1. \)

Other player cheats: ensure \( O(\sqrt{n}) \) regret after \( n \) rounds.
Both players honest: \(-\minimax\) equilibria at rate \(-1\).

Other player cheats: ensure \(O(\sqrt{n})\) regret after \(n\) rounds.
Zero-sum Games with Uncoupled Dynamics

- Both players honest: \( -\minimax \) equilibrium at rate \( -1 \).
- Other player cheats: ensure \( O(\sqrt{n}) \) regret after \( n \) rounds.
Zero-sum Games with Uncoupled Dynamics

Both players honest: \( -\text{minimax equilibria at rate } \varepsilon - 1 \).

Other player cheats: ensure \( O(\sqrt{n}) \) regret after \( n \) rounds.

\[ p_t \in \Delta_N \]

\[ q_t \in \Delta_M \]

\[ A \]
Both players honest: $\epsilon$-minimax equilibrium at rate $\epsilon^{-1}$. 
Both players honest: $\epsilon$-minimax equilibria at rate $\epsilon^{-1}$. 
Both players honest: $\varepsilon$-minimax equilibria at rate $\varepsilon^{-1}$.

Other player cheats: ensure $O(\sqrt{n})$ regret after $n$ rounds.
For $t = 1$ to $n$
    Play distribution $q_t \in \Delta_M$
    Suffer loss $q_t^T A p_t$ and observe $A q_t$
End
For \( t = 1 \) to \( n \)

Play distribution \( q_t \in \Delta_M \)

Suffer loss \( q_t^\top A p_t \) and observe \( A q_t \)

End

**When both players are honest:** Neither \( A \), nor \( N \) is known.

We want strategy for the players so that:

\[
\left| \frac{1}{n} \sum_{t=1}^{n} q_t^\top A p_t - \min_{q \in \Delta_M} \max_{p \in \Delta_N} q^\top A p \right| \leq O(n^{-1})
\]
For $t = 1$ to $n$

- Play distribution $q_t \in \Delta_M$
- Suffer loss $q_t^T Ap_t$ and observe $Aq_t$

End

**When both players are honest:** Neither $A$, nor $N$ is known.

We want strategy for the players so that:

$$\left| \frac{1}{n} \sum_{t=1}^{n} q_t^T Ap_t - \min_{q \in \Delta_M} \max_{p \in \Delta_N} q^T Ap \right| \leq O(n^{-1})$$

**When other player is dishonest:**

$$\frac{1}{n} \sum_{t=1}^{n} q_t^T Ap_t - \inf_{i \in [M]} \sum_{t=1}^{n} e_i^T Ap_t \leq O(n^{-1/2})$$
If both player play exponential weights, we get regret bound of $\sqrt{n}$ and convergence to equilibria at $n^{-1/2}$ (when honest)

Can we use adaptive learning algorithms for both players to get faster rate when players are honest?
Mirror Descent With Predictable Sequence

\( \hat{Y} = \mathcal{F} \) convex subset of vector space (unit ball under norm \( \| \cdot \| \))
Mirror Descent With Predictable Sequence

- $\hat{Y} = \mathcal{F}$ convex subset of vector space (unit ball under norm $\|\cdot\|$)
- For each $y \in \mathcal{Y}$, $\ell(\cdot, y)$ is convex and 1-Lipschitz w.r.t. $\|\cdot\|$. 

$\hat{y}_t = \arg\min_{\hat{y} \in \mathcal{F}} \ell(\hat{y}, \mathcal{M}_t) + R(\hat{y}, z_{t-1})$

$z_t = \arg\min_{\hat{y} \in \mathcal{F}} \ell(\hat{y}, y_t) + \nabla \ell(\hat{y}_t, y_t) + R(\hat{y}, z_{t-1})$
Mirror Descent With Predictable Sequence

- $\hat{Y} = \mathcal{F}$ convex subset of vector space (unit ball under norm $\|\cdot\|$)
- For each $y \in \mathcal{Y}$, $\ell(\cdot, y)$ is convex and 1-Lipschitz w.r.t. $\|\cdot\|$.
- Let $(M_t)_{t \geq 1}$ be any predictable sequence (computable at round $t$)
Mirror Descent With Predictable Sequence

- \( \hat{Y} = \mathcal{F} \) convex subset of vector space (unit ball under norm \( \| \cdot \| \))
- For each \( y \in Y \), \( \ell(\cdot, y) \) is convex and 1-Lipschitz w.r.t. \( \| \cdot \| \).
- Let \( (M_t)_{t \geq 1} \) be any predictable sequence (computable at round \( t \))
- Mirror descent with predictable sequence:

\[
\hat{y}_t = \arg\min_{\hat{y} \in \mathcal{F}} \eta_t \langle \hat{y}, M_t \rangle + \Delta_{\mathcal{R}}(\hat{y}|z_{t-1}) , \quad z_t = \arg\min_{\hat{y} \in \mathcal{F}} \eta_t \langle \hat{y}, \nabla \ell(\hat{y}_t, y_t) \rangle + \Delta_{\mathcal{R}}(\hat{y}|z_{t-1})
\]

\( \mathcal{R} \) is 1-strongly convex w.r.t. \( \| \cdot \| \) and \( \Delta_{\mathcal{R}} \) is Bregman divergence.
Mirror Descent With Predictable Sequence

- \( \hat{Y} = \mathcal{F} \) convex subset of vector space (unit ball under norm \( \| \cdot \| \))
- For each \( y \in \mathcal{Y} \), \( \ell(\cdot, y) \) is convex and 1-Lipschitz w.r.t. \( \| \cdot \| \).
- Let \( (M_t)_{t \geq 1} \) be any predictable sequence (computable at round \( t \))
- Mirror descent with predictable sequence:

\[
\hat{y}_t = \arg\min_{\hat{y} \in \mathcal{F}} \eta_t \langle \hat{y}, M_t \rangle + \Delta_\mathcal{R}(\hat{y} | z_{t-1}), \quad z_t = \arg\min_{\hat{y} \in \mathcal{F}} \eta_t \langle \hat{y}, \nabla \ell(\hat{y}_t, y_t) \rangle + \Delta_\mathcal{R}(\hat{y} | z_{t-1})
\]

\( \mathcal{R} \) is 1-strongly convex w.r.t. \( \| \cdot \| \) and \( \Delta_\mathcal{R} \) is Bregman divergence.

- Below adaptive bound is achievable with appropriate \( \eta_t \)

\[
\text{Reg}_n \leq O\left( \frac{R_{\max}}{2} \sqrt{\sum_{t=1}^{n} \| \nabla_t - M_t \|_*^2} - C \sum_{t=1}^{n} \| \hat{y}_t - z_{t-1} \|^2 \right)
\]
\[
\hat{Y} = \mathcal{F} = \Delta_M, \mathcal{V} = \Delta_n, \ell(q, p) = q^\top A p
\]

Assume entries of \( A \) are bounded by 1.

Use \( M_t = A p_{t-1}, \mathcal{R}(q) = \sum_{i=1}^{M} q_i \log(q_i) \).
**Back to Uncoupled Dynamics Game**

- $\hat{Y} = \mathcal{F} = \Delta_M, \mathcal{Y} = \Delta_n, \ell(q, p) = q^\top Ap$
- Assume entries of $A$ are bounded by 1.
- Use $M_t = Ap_{t-1}$, $\mathcal{R}(q) = \sum_{i=1}^{M} q_i \log(q_i)$.
- Regret for player I:

$$\text{Reg}_n^I \leq O\left( \log M \sqrt{\sum_{t=1}^{n} \|p_t - p_{t-1}\|_1^2 - C \sum_{t=1}^{n} \|q_t - q_{t-1}\|_1^2} \right) = O(\log M\sqrt{n})$$
\[ \hat{Y} = \mathcal{F} = \Delta_M, \mathcal{Y} = \Delta_n, \ell(q, p) = q^\top A p \]

- Assume entries of \( A \) are bounded by 1.
- Use \( M_t = A p_{t-1}, \mathcal{R}(q) = \sum_{i=1}^{M} q_i \log(q_i) \).
- Regret for player I:

\[
\text{Reg}_n^I \leq O \left( \log M \sqrt{\sum_{t=1}^{n} \|p_t - p_{t-1}\|_1^2} - C \sum_{t=1}^{n} \|q_t - q_{t-1}\|_1^2 \right) = O(\log M\sqrt{n})
\]

- Convergence to Equilibrium:

\[
\left| \frac{1}{n} \sum_{t=1}^{n} q_t^\top A p_t - \min_{q \in \Delta_M} \max_{p \in \Delta_N} q^\top A p \right| \leq \frac{1}{n} \left( \text{Reg}_n^I + \text{Reg}_n^{II} \right) = O \left( \frac{\log(NMn)}{n} \right)
\]
In fact if other player uses exponential weights (or other OL algorithm) with step size $\eta_t$ at time $t$, regret of first player improves to

$$\text{Reg}^I_n \leq O\left(\sqrt{\sum_{t=1}^{n} \eta_t^2}\right)$$
In fact if other player uses exponential weights (or other OL algorithm) with step size $\eta_t$ at time $t$, regret of first player improves to

$$\text{Reg}_n^I \leq O\left(\sqrt{\sum_{t=1}^{n} \eta_t^2}\right)$$

Can be extended to case when feedback is only expected loss $q_t^\top A p_t$
In fact if other player uses exponential weights (or other OL algorithm) with step size $\eta_t$ at time $t$, regret of first player improves to

$$\text{Reg}^I_n \leq O\left(\sqrt{\sum_{t=1}^{n} \eta_t^2}\right)$$

Can be extended to case when feedback is only expected loss $q_t^\top Ap_t$

[Syrgkanis et al’15] Extend the results to multiplayer games and correlated equilibria
In fact if other player uses exponential weights (or other OL algorithm) with step size $\eta_t$ at time $t$, regret of first player improves to

$$\text{Reg}^I_n \leq O\left(\sqrt{\sum_{t=1}^{n} \eta_t^2}\right)$$

Can be extended to case when feedback is only expected loss $q_t^\top A p_t$

[Syrgkanis et al’15] Extend the results to multiplayer games and correlated equilibria

Main message: Adaptive regret bounds can harness dynamics of other player(s) to converge faster
Adaptive Online Learning and Optimization
\[
\operatorname{argmax}_{x \in \mathcal{X}} c^\top x \quad \text{s.t. } \forall i \in [d], \ G_i(x) \leq 1
\]

Each \( G_i \) is convex and smooth.
Approximate Convex Programming

\[
\arg\max_{x \in X} c^\top x \quad \text{s.t. } \forall i \in [d], \ G_i(x) \leq 1
\]

Each \( G_i \) is convex and smooth.

- Assume we know value at optimal (if not do binary search) and add this as linear constraint
Approximate Convex Programming

\[
\arg\max_{x \in X} c^\top x \quad \text{s.t.} \quad \forall i \in [d], \quad G_i(x) \leq 1
\]

Each \( G_i \) is convex and smooth.

- Assume we know value at optimal (if not do binary search) and add this as linear constraint
- Treat the optimization problem as two player game

Use gradient descent with predictable sequence for player I

Use exponential weights with predictable sequence for player II

Compared to typical regret minimizing approach, we get \( 1/\varepsilon \) rate.
\[ \argmax_{x \in X} c^\top x \quad \text{s.t. } \forall i \in [d], \ G_i(x) \leq 1 \]

Each $G_i$ is convex and smooth.

- Assume we know value at optimal (if not do binary search) and add this as linear constraint
- Treat the optimization problem as two player game
- Player one plays $x_t \in X$ and player two plays mixed constraints
argmax \ c^\top x \quad \text{s.t. } \forall i \in [d], \ G_i(x) \leq 1

Each \( G_i \) is convex and smooth.

- Assume we know value at optimal (if not do binary search) and add this as linear constraint
- Treat the optimization problem as two player game
- Player one plays \( x_t \in \mathcal{X} \) and player two plays mixed constraints
- Use gradient descent with predictable sequence for player I
Approximate Convex Programming

\[
\arg\max_{x \in X} \ c^\top x \quad \text{s.t.} \quad \forall i \in [d], \ G_i(x) \leq 1
\]

Each \( G_i \) is convex and smooth.

- Assume we know value at optimal (if not do binary search) and add this as linear constraint
- Treat the optimization problem as two player game
- Player one plays \( x_t \in X \) and player two plays mixed constraints
- Use gradient descent with predictable sequence for player I
- Use exponential weights with predictable sequence for player II
**Approximate Convex Programming**

$$\arg\max_{x \in \mathcal{X}} c^\top x \quad \text{s.t.} \quad \forall i \in [d], \quad G_i(x) \leq 1$$

Each $G_i$ is convex and smooth.

- Assume we know value at optimal (if not do binary search) and add this as linear constraint
- Treat the optimization problem as two player game
- Player one plays $x_t \in \mathcal{X}$ and player two plays mixed constraints
- Use gradient descent with predictable sequence for player I
- Use exponential weights with predictable sequence for player II
- Compared to typical regret minimizing approach, we get $1/\epsilon$ rate.
Approximate Convex Programming

\[
\arg\max_{x \in \mathcal{X}} c^\top x \quad \text{s.t. } \forall i \in [d], \ G_i(x) \leq 1
\]

Each \(G_i\) is convex and smooth.

Lemma

If there exists \(f_0 \in \mathcal{F}\) s.t. \(G_i(f_0) \leq 1 - \gamma\) and \(c^\top f_0 \geq 0\), then in time \(\frac{d B \gamma \sqrt{\log d}}{\varepsilon}\), we can find solution \(\hat{x} \in \mathcal{X}\) such that, \(\forall i \in [d], G_i(\hat{x}) \leq 1\),

\[
c^\top \hat{x} \geq (1 - \varepsilon) \sup_{x \in \mathcal{X} : \forall i, G_i(x) \leq 1} c^\top x
\]
Adaptive Online Learning and Deviation Bounds
Let \((X_t)_{t \geq 1}\) be a m.d.s. taking values in \((2, D)\)-smooth Banach space with norm \(|\cdot|\) s.t., \(\sum_t X_t^2 \leq \sigma\). By [Pinelis’94],

\[
P \left( \sup_{n \geq 0} \left\| \sum_{t=1}^{n} X_t \right\| > \sigma t \right) \leq \exp \left( -\frac{t^2}{D^2} \right)
\]
Let \((X_t)_{t \geq 1}\) be a m.d.s. taking values in \((2, D)\)-smooth Banach space with norm \(|\cdot|\) s.t., \(\sum_t \|X_t\|_\infty^2 \leq \sigma\). By [Pinelis’94],

\[
P\left(\sup_{n \geq 0} \left\| \sum_{t=1}^{n} X_t \right\| > \sigma t \right) \leq \exp\left(-\frac{t^2}{D^2}\right)
\]

Can we replace \(\sigma\) by a distribution dependent quantity?
Let \((X_t)_{t \geq 1}\) be a m.d.s. taking values in \((2, D)\)-smooth Banach space with norm \(\| \cdot \|\) s.t., \(\sum_t \|X_t\|_\infty^2 \leq \sigma\). By [Pinelis’94],

\[
P \left( \sup_{n \geq 0} \left\| \sum_{t=1}^{n} X_t \right\| > \sigma t \right) \leq \exp \left( -\frac{t^2}{D^2} \right)
\]

Can we replace \(\sigma\) by a distribution dependent quantity?

Can we prove the bound easily maybe even just for Euclidean norm?
Consider online linear optimization with $\mathbf{F} = \mathbf{B} \cdot \mathbf{B}^\ast$, we get adaptive regret bound for any $\mathbf{X}_1, \ldots, \mathbf{X}_n$.

If adversary played $\mathbf{m.d.s.}$, then $\hat{\mathbf{y}}_t, - \mathbf{X}_t$ is a real valued $\mathbf{m.d.s.}$ with magnitude bounded by $\mathbf{X}_t$.

Rewriting:

$$\left\| \hat{\mathbf{y}}_t, - \mathbf{X}_t \right\| \leq \sqrt{2e^{-\frac{1}{16}}}$$
Consider online linear optimization with $F = B_{\|\cdot\|_*}$, we get adaptive regret bound for any $X_1, \ldots, X_n$,

$$\sum_{t=1}^{n} \langle \hat{y}_t, X_t \rangle - \inf_{f: \|f\|_* \leq 1} \sum_{t=1}^{n} \langle f, X_t \rangle \leq O \left( D \sqrt{\sum_{t=1}^{n} \| X_t \|^2} \right)$$
Consider online linear optimization with $\mathcal{F} = B\|\cdot\|_*$, we get adaptive regret bound for any $X_1, \ldots, X_n$,

$$\sum_{t=1}^{n} \langle \hat{y}_t, X_t \rangle - \inf_{f: \|f\|_* \leq 1} \sum_{t=1}^{n} \langle f, X_t \rangle \leq O \left( D \sqrt{\sum_{t=1}^{n} \|X_t\|^2} \right)$$

**Rewriting:**

$$\left\| \sum_{t=1}^{n} X_t \right\| - D \sqrt{\sum_{t=1}^{n} \|X_t\|^2} \leq \sum_{t=1}^{n} \langle \hat{y}_t, -X_t \rangle$$
Consider online linear optimization with $\mathcal{F} = B_{\|\cdot\|_*}$, we get adaptive regret bound for any $X_1, \ldots, X_n$,

$$
\sum_{t=1}^{n} \langle \hat{y}_t, X_t \rangle - \inf_{f: \|f\|_* \leq 1} \sum_{t=1}^{n} \langle f, X_t \rangle \leq O \left( D \sqrt{\sum_{t=1}^{n} \|X_t\|^2} \right)
$$

Rewriting:

$$
\left\| \sum_{t=1}^{n} X_t \right\| - D \sqrt{\sum_{t=1}^{n} \|X_t\|^2} \leq \sum_{t=1}^{n} \langle \hat{y}_t, -X_t \rangle
$$

If adversary played m.d.s. $X_t$’s, then $\langle \hat{y}_t, -X_t \rangle$ is a real valued m.d.s. with magnitude bounded by $\|X_t\|$.

$$
P \left( \left\| \sum_{t=1}^{n} X_t \right\| - D \sqrt{\sum_{t=1}^{n} \|X_t\|^2} > t \Sigma \right) \leq P \left( \sum_{t=1}^{n} \langle \hat{y}_t, -X_t \rangle > t \Sigma \right) \leq \sqrt{2} e^{-\frac{t^2}{16}}
$$
Can replace $\sigma$ by

$$
\Sigma = \sqrt{\sum_{t=1}^{n}(\|X_t\|^2 + \mathbb{E}_{t-1}[\|X_t\|^2]) + \left(\mathbb{E}\left[\sqrt{\sum_{t=1}^{n}(\|X_t\|^2 + \mathbb{E}_{t-1}[\|X_t\|^2])}\right]\right)^2}
$$
Can replace $\sigma$ by

$$
\Sigma = \sqrt{\sum_{t=1}^{n}(\|X_t\|^2 + \mathbb{E}_{t-1}[\|X_t\|^2]) + \left(\mathbb{E}\left[\sqrt{\sum_{t=1}^{n}(\|X_t\|^2 + \mathbb{E}_{t-1}[\|X_t\|^2])}\right]\right)^2}
$$

On the other hand, assume that the deviation inequality holds, then integrating the tails, for any m.d.s.,

$$
\mathbb{E}\left[\left\|\sum_{t=1}^{n}X_t\right\| - K\sqrt{\sum_{t=1}^{n}\|X_t\|^2}\right] \leq 0
$$
**Deviation Inequalities: Warmup**

- Can replace $\sigma$ by

\[
\Sigma = \sqrt{\sum_{t=1}^{n} (\|X_t\|^2 + \mathbb{E}_{t-1} [\|X_t\|^2])} + \left( \mathbb{E} \left[ \sqrt{\sum_{t=1}^{n} (\|X_t\|^2 + \mathbb{E}_{t-1} [\|X_t\|^2])} \right] \right)^2
\]

- On the other hand, assume that the deviation inequality holds, then integrating the tails, for any m.d.s.,

\[
\mathbb{E} \left[ \left\| \sum_{t=1}^{n} X_t \right\| - K \sqrt{\sum_{t=1}^{n} \|X_t\|^2} \right] \leq 0
\]

- However the above implies the regret bound: (using minimax theorem repeatedly and moving to dual game and overbounding)

\[
\sum_{t=1}^{n} \langle \hat{y}_t, X_t \rangle - \inf_{f : \|f\|_* \leq 1} \sum_{t=1}^{n} \langle f, X_t \rangle \leq O \left( \sqrt[2]{\sum_{t=1}^{n} \|X_t\|^2} \right)
\]
Adaptive regret bound for OLO + Hoeffding Azuma (++) enough to prove concentration in Banach space

One-one correspondence between regret statement and deviation inequalities
Adaptive regret bound for OLO + Hoeffding Azuma (++) enough to prove concentration in Banach space

One-one correspondence between regret statement and deviation inequalities

Is this phenomenon true in general?
Consider supervised learning game with linear loss:

For $t = 1$ to $n$

- Adversary picks $x_t \in \mathcal{X}$
- Learner picks $\hat{y}_t \in \mathbb{R}$
- Adversary picks $y_t \in \{\pm 1\}$
- Learner suffers loss $\hat{y}_t \cdot y_t$
Rough statement: we say \( \mathcal{F} \) was bounded by 1.
DEVIATION INEQUALITIES

Rough statement: say $\mathcal{F}$ was bounded by 1,

1 Adaptive regret bound

\[
\sum_{t=1}^{n} \hat{y}_t \cdot y_t \leq \inf_{f \in \mathcal{F}} \left\{ \sum_{t=1}^{n} y_t f(x_t) + B(x_1, \ldots, x_n; f) \right\} \text{ implies,}
\]
Rough statement: say \( \mathcal{F} \) was bounded by 1,

1. **Adaptive regret bound**

\[
\sum_{t=1}^{n} \hat{y}_t \cdot y_t \leq \inf_{f \in \mathcal{F}} \left\{ \sum_{t=1}^{n} y_t f(x_t) + B(x_1, \ldots, x_n; f) \right\} \quad \text{implies,}
\]

2. **High probability bound**

\[
P \left( \sup_{f \in \mathcal{F}} \left\{ \sum_{t=1}^{n} f(X_t) - \mathbb{E}_{t-1} [f(X_t)] - B(X_1, \ldots, X_n; f) \right\} > t \right) \leq e^{-\frac{t^2}{n}} \quad \text{implies,}
\]
Deviation Inequalities

Rough statement: \( \mathcal{F} \) was bounded by 1,

1. Adaptive regret bound
\[
\sum_{t=1}^{n} \hat{y}_t \cdot y_t \leq \inf_{f \in \mathcal{F}} \left\{ \sum_{t=1}^{n} y_t f(x_t) + B(x_1, \ldots, x_n; f) \right\} \quad \text{implies,}
\]

2. High probability bound
\[
P\left( \sup_{f \in \mathcal{F}} \left\{ \sum_{t=1}^{n} f(X_t) - \mathbb{E}_{t-1} [f(X_t)] - B(X_1, \ldots, X_n; f) \right\} > t \right) \leq e^{-\frac{t^2}{n}} \quad \text{implies,}
\]

3. Bound on Expectation
\[
\mathbb{E} \left[ \sup_{f \in \mathcal{F}} \left\{ \sum_{t=1}^{n} f(X_t) - \mathbb{E}_{t-1} [f(X_t)] - B(X_1, \ldots, X_n; f) \right\} \right] \leq \sqrt{n}
\]

Above implies 1 with worse \( B \)
Proving regret bounds and proving high-probability tail bounds for the supremum of a collection of martingales are equivalent.
**Deviation Bounds**

- Proving regret bounds and proving high-probability tail bounds for the supremum of a collection of martingales are equivalent.
- Instead of linear game using game with squared loss yields tighter concentration for less complex $\mathcal{F}$.

Example: $Z_1, \ldots, Z_n \in \mathbb{R}^d$ are m.d.s.

$$
\mathbb{E} \left[ \max_{j \leq d} \left\{ \left| \sum_{t=1}^{n} Z_t[j] \right| - \sqrt{2 \log(d) \sum_{t=1}^{n} Z_t^2[j]} \right\} \right] \leq 0
$$

"Each function only pays its variance"
Deviation Bounds

- Proving regret bounds and proving high-probability tail bounds for the supremum of a collection of martingales are equivalent.
- Instead of linear game using game with squared loss yields tighter concentration for less complex $\mathcal{F}$.

Example: $Z_1, \ldots, Z_n \in \mathbb{R}^d$ are m.d.s.

$$\mathbb{E} \left[ \max_{j \leq d} \left\{ \left| \sum_{t=1}^{n} Z_t[j] \right| - \sqrt{2 \log(d) \sum_{t=1}^{n} Z_t^2[j]} \right\} \right] \leq 0$$

- “Each function only pays its variance”
Deviation Bound

- If $\mathcal{F}$ has Seq. Entropy $(\mathcal{F}, \alpha) \sim \alpha^{-q}$, then,
  - If $q \geq 2$, then for $p = q/(q - 1)$
    \[
    P \left( \sup_{f \in \mathcal{F}} \sum_{t=1}^{n} (f(X_t) - \mathbb{E}_{t-1}[f(X_t)]) - n^{1/p}(1 + t) \right) \leq C \exp(-ct^2)
    \]
  - If $q < 2$, then
    \[
    P \left( \sup_{f \in \mathcal{F}} \left\{ \sum_{t=1}^{n} (f(X_t) - \mathbb{E}_{t-1}[f(X_t)]) - n^{q/4} \text{Var}(f)^{2/q} - t\sqrt{\text{Var}(f)} \right\} > 0 \right) \leq C \exp(-ct^2)
    \]
    where $\text{Var}(f) = \sum_{t=1}^{n} (f(X_t) - \mathbb{E}_{t-1}[f(X_t)])^2$
Adaptive online learning algorithms in strategies for games that can harness properties of adversaries strategies

Adaptive online learning algorithms for approximate convex programming with better dependence on $\epsilon$

Regret minimization equivalent to deviation inequalities for martingales
Adaptive online learning algorithms in strategies for games that can harness properties of adversaries strategies

Adaptive online learning algorithms for approximate convex programming with better dependence on $\epsilon$

Regret minimization equivalent to deviation inequalities for martingales

Some extensions:

- Extend notion of martingale type beyond Banach spaces to arbitrary $\mathcal{F}$
- Finer control of expectation of supremum via per-function variance for simpler $\mathcal{F}$