Last Lecture

Online Vs Statistical

- Often, online and statistical learning rates match.
- $\mathcal{R}_n^{sq}(\ell \circ \mathcal{F}) \leq C\, \mathcal{R}_n^{stat}(\ell \circ \mathcal{F})$.

(When) Can we use this observation to obtain faster algorithms?

Modifying Rademacher Relaxation

- Earlier we went for the rate given by

$$\phi((x_1, y_1), \ldots, (x_n, y_n)) = -\inf_{f \in \mathcal{F}} \sum_{t=1}^{n} \ell(f(x_t), y_t) + 2\, \mathcal{R}_n^{sq}(\ell \circ \mathcal{F})$$

- Using previous observation let us weaken it first to

$$\phi^{rand}((x_1, y_1), \ldots, (x_n, y_n)) = -\inf_{f \in \mathcal{F}} \sum_{t=1}^{n} \ell(f(x_t), y_t) + 2C\, \mathcal{R}_n^{stat}(\ell \circ \mathcal{F})$$

- Now the key to cutting down computation is to replace the $\mathbf{x}, \mathbf{y}$ trees by draws of future instances from some fixed distribution D

MODIFYING RADEMACHER RELAXATION

- To this end, let us guess a Rademacher relaxation as:

$$\mathbf{Rad}_n^{rand}(x_{1:t}, y_{1:t})$$
$$= \mathop{\mathbb{E}}_{(x,y)_{t+1:n} \sim D} \mathbb{E}_\epsilon \sup_{f \in \mathcal{F}} \left[2C \sum_{s=t+1}^{n} \epsilon_s \ell(f(x_s), y_s) - \sum_{s=1}^{t} \ell(f(x_s), y_s) \right] - 2C\, \mathcal{R}_n^{stat}(\ell \circ \mathcal{F})$$

- Clearly $\mathbf{Rad}_n^{rand}(x_{1:n}, y_{1:n}) = \phi^{rand}((x_1, y_1), \ldots, (x_n, y_n))$

- Also,

$$\mathbf{Rad}_n^{rand}(\cdot) = 2C\, \mathbb{E}_{S \sim D} \widehat{\mathcal{R}}_n(\ell \circ \mathcal{F}) - 2C\, \mathcal{R}_n^{stat}(\ell \circ \mathcal{F}) \leq 0$$

- So dominance and final value condition are satisfied.

Modifying Rademacher Relaxation

- Coming to Admissibility: we need to show that For any $x_t \in \mathcal{X}$

$$\inf_{q \in \hat{\mathcal{Y}}} \sup_{y} \left\{ \mathbb{E}_{\hat{y}_t \sim q} \ell(\hat{y}_t, y_t) + \mathbf{Rad}_n^{rand}(x_{1:t}, y_{1:t}) \right\} \leq \mathbf{Rad}_n^{rand}(x_{1:t-1}, y_{1:t-1})$$

- In general this condition need not be satisfies. But we will show that the following condition suffices.

- Condition: For all t, $\forall \epsilon_{t+1:n} \in \{\pm 1\}^{n-t}$, $\forall x_{1:t-1,t+1:n}, y_{1:t-1,t+1:n}$,

$$\sup_{x_t, y_t} \mathbb{E}_{\epsilon_n} \left[\sup_{f \in \mathcal{F}} \left[2C \sum_{s=t+1}^{n} \epsilon_s \ell(f(x_s), y_s) + 2\epsilon_t \ell(f(x_t), y_t) - \sum_{s=1}^{t-1} \ell(f(x_s), y_s) \right] \right.$$

$$\leq \mathbb{E}_{(x_t, y_t) \sim D} \mathbb{E}_{\epsilon_n} \left[\sup_{f \in \mathcal{F}} \left[2C \sum_{s=t}^{n} \epsilon_s \ell(f(x_s), y_s) - \sum_{s=1}^{t-1} \ell(f(x_s), y_s) \right] \right.$$

Random Playout

Define $R_t = x_{t+1:n}, y_{t+1:n}, \epsilon_{t+1:n}$ and let $D_t = D^{n-t} \times \text{Unif}\{\pm 1\}^{n-t}$

$$\phi_t(x_{1:t}, y_{1:t}; R_t) = \sup_{f \in \mathcal{F}} \left\{ 2C \sum_{s=t+1}^{n} \epsilon_s \ell(f(x_s), y_s) - \sum_{s=1}^{t} \ell(f(x_s), y_s) \right\}$$

Algorithm : **Draw** $R_t \sim D^t$**, and return,**

$$\tilde{q}_t(R_t) = \operatorname*{argmin}_{q \in \Delta(\mathcal{Y})} \sup_{y_t} \left\{ \mathbb{E}_{\hat{y}_t \sim q_t} \left[\ell(\hat{y}_t, y_t) \right] + \Phi_t(x_{1:t}, y_{1:t}, R_t) \right\}$$

Why/When does this work?

Random Playout

To show admissibility

$$
\begin{aligned}
&\inf_{q_t}\sup_{y_t}\left\{\mathbb{E}_{\hat{y}_t\sim q_t}\left[\ell(\hat{y}_t,y_t)\right]+\mathbf{Rel}_n\left(x_{1:t},y_{1:t}\right)\right\}\\
&=\inf_{q_t}\sup_{y_t}\left\{\mathbb{E}_{\hat{y}_t\sim q_t}\left[\ell(\hat{y}_t,y_t)\right]+\mathbb{E}_{R_t\sim D^t}\left[\Phi_t(x_{1:t},y_{1:t},R_t)\right]\right\}\\
&\le\sup_{y_t}\left\{\mathbb{E}_{\hat{y}_t\sim\tilde{q}_t}\left[\ell(\hat{y}_t,y_t)\right]+\mathbb{E}_{R_t\sim D^t}\left[\Phi_t(x_{1:t},y_{1:t},R_t)\right]\right\}\\
&=\sup_{y_t}\left\{\mathbb{E}_{R_t\sim D^t}\left[\mathbb{E}_{\hat{y}_t\sim\tilde{q}_t(R_t)}\left[\ell(\hat{y}_t,y_t)\right]\right]+\mathbb{E}_{R_t\sim D_t}\left[\Phi_t(x_{1:t},y_{1:t},R_t)\right]\right\}\\
&\le\mathbb{E}_{R_t\sim D^t}\left[\sup_{y_t}\left\{\mathbb{E}_{\hat{y}_t\sim\tilde{q}_t(R_t)}\left[\ell(\hat{y}_t,y_t)\right]+\Phi_t(x_{1:t},y_{1:t},R_t)\right\}\right]\\
&=\mathbb{E}_{R_t\sim D^t}\left[\inf_{q_t}\sup_{y_t}\left\{\mathbb{E}_{\hat{y}_t\sim q_t}\left[\ell(\hat{y}_t,y_t)\right]+\Phi_t(x_{1:t},y_{1:t},R_t)\right\}\right]\\
&=\mathbb{E}_{R_t\sim D^t}\left[\sup_{p_t}\left\{\inf_{\hat{y}_t\in\mathcal{Y}}\mathbb{E}_{y_t\sim p_t}\left[\ell(\hat{y}_t,y_t)\right]+\mathbb{E}_{y_t\sim p_t}\left[\Phi_t(x_{1:t},y_{1:t},R_t)\right]\right\}\right]
\end{aligned}
$$

Random Playout

To finish admissibility, note that

$$\sup_{p_t}\left\{\inf_{\hat{y}_t\in\mathcal{Y}}\mathbb{E}_{y_t\sim p_t}\left[\ell(\hat{y}_t,y_t)\right]+\mathbb{E}_{y_t\sim p_t}\left[\Phi_t(x_{1:t},y_{1:t},R_t)\right]\right\}$$

$$=\sup_{p_t}\left\{\inf_{\hat{y}_t\in\mathcal{Y}}\mathbb{E}_{y_t\sim p_t}\left[\ell(\hat{y}_t,y_t)\right]+\mathbb{E}_{y_t\sim p_t}\left[\sup_{f\in\mathcal{F}}\left\{2C\sum_{s=t+1}^{n}\epsilon_s\ell(f(x_s),y_s)-\sum_{s=1}^{t}\ell(f(x_s),y_s)\right\}\right]\right\}$$

$$\le\sup_{x_t}\mathbb{E}_{\epsilon_t}\left[\sup_{f\in\mathcal{F}}\left\{2C\sum_{s=t+1}^{n}\epsilon_s\ell(f(x_s),y_s)+2\epsilon_t\ell(f(x_t),y_t)-\sum_{s=1}^{t-1}\ell(f(x_s),y_s)\right\}\right]$$

But our sufficient condition was:

$$\sup_{x_t,y_t}\mathbb{E}_{\epsilon_n}\left[\sup_{f\in\mathcal{F}}\left[2C\sum_{s=t+1}^{n}\epsilon_s\ell(f(x_s),y_s)+2\epsilon_t\ell(f(x_t),y_t)-\sum_{s=1}^{t-1}\ell(f(x_s),y_s)\right]\right]$$

$$\le\mathbb{E}_{(x_t,y_t)\sim D}\mathbb{E}_{\epsilon_n}\left[\sup_{f\in\mathcal{F}}\left[2C\sum_{s=t}^{n}\epsilon_s\ell(f(x_s),y_s)-\sum_{s=1}^{t-1}\ell(f(x_s),y_s)\right]\right]$$

Hence,

$$\inf_{q_t}\sup_{y_t}\left\{\mathbb{E}_{\hat{y}_t\sim q_t}\left[\ell(\hat{y}_t,y_t)\right]+\mathbf{Rel}_n\left(x_{1:t},y_{1:t}\right)\right\}$$

$$\leq \mathbb{E}_{(x,y)_{t+1:n}\sim D^{n-t}}\mathbb{E}_{\epsilon_{t+1:n}}\mathbb{E}_{(x_t,y_t)\sim D}\mathbb{E}_{\epsilon_n}\left[\sup_{f\in\mathcal{F}}\left[2C\sum_{s=t}^{n}\epsilon_s\ell(f(x_s),y_s)-\sum_{s=1}^{t-1}\ell(f(x_s),y_s)\right]\right]$$

$$=\mathbf{Rel}_n\left(x_{1:t-1},y_{1:t-1}\right)$$

Example: Linear Predictors

- Online linear optimization, $\mathcal{F} = \{f : \|f\| \leq 1\}$, $\mathbf{D} = \{\nabla : \|\nabla\|_* \leq 1\}$
- Condition: $\exists D$ and constant C, such that, for any vector w,

$$\sup_{x_t} \mathbb{E}_{\epsilon_t} \left[\|w + 2\epsilon_t x_t\|_* \right] \leq \mathbb{E}_{x_t \sim D} \left[\|w + C x_t\|_* \right]$$

- ℓ_1^d / ℓ_∞^d : $D = \text{Unif}\{\pm 1\}^d$ or any other symmetric distribution on each coordinate (Eg. normal distribution)
- Algorithm : Round t draw $R_t \sim N(0, (n-t) I_d)$

$$\hat{y}_t = \underset{i \in [d]}{\text{argmin}} \left| \sum_{j=1}^{t-1} \nabla_j[i] + R_t[i] \right|$$

- Bound : $\mathbb{E}\left[\text{Reg}_n\right] \leq \frac{1}{n} \mathbf{Rel}_n(\cdot) = O\left(\sqrt{\frac{\log d}{n}}\right)$

Rough Sketch of Proof

- $w = 2C\sum_{s=t+1}^{n} \nabla_s - \sum_{s=1}^{t-1} \nabla_s$ where $\nabla_{1:t-1}$ are past losses and $\nabla_{t+1:n}$ are drawn from $\text{Unif}\{-1,1\}^d$
- Assume $t < n - \sqrt{n}$, for last $\sqrt{n}$ rounds even if we are completely off, regret bound does not change
- Hence w can be seen as vector $-\sum_{s=1}^{t-1} \nabla_s$ where each coordinate is perturbed by $2C\sum_{s=t+1}^{n} \nabla_s$
- With very high probability, if i^* and j^* are top two coordinates of w, $|w[i^*]| - |w[j^*]| > 4$, hence, with high probability,

$$\sup_{x_t \in [-1,1]^d} \mathbb{E}_{\epsilon_t}\left[\|w + 2\epsilon_t x_t\|_\infty\right] = \sup_{x_t \in [-1,1]^d} \mathbb{E}_{\epsilon_t}\left[|w[i^*] + 2\epsilon_t x_t[i^*]|\right]$$
$$= \mathbb{E}_{\epsilon_t}\left[|w[i^*] + 2\epsilon_t|\right] = \mathbb{E}_{x_t \sim D}\left[\|w + 2\epsilon_t x_t\|_\infty\right]$$

- In general we don't need this high probability stuff, we can directly prove the condition, just need to check cases.

Rough Sketch of Proof

- Why update of form $\hat{y}_t = \text{argmin}_{i\in[d]} \left|\sum_{j=1}^{t} \nabla_j[i] + R_t[i]\right|$
- To see this, note that the algorithm we need is originally of form,

$$\begin{aligned}
\hat{y}_t &= \underset{\hat{y}\in\mathcal{F}}{\text{argmin}} \sup_{\nabla_t} \left\{ \langle \hat{y}, \nabla_t \rangle + \sup_{f\in\mathcal{F}} \left\{ \langle f, -R_t \rangle - \left\langle f, \sum_{s=1}^{t} \nabla_s \right\rangle \right\} \right\} \\
&= \underset{\hat{y}\in\mathcal{F}}{\text{argmin}} \sup_{f\in\mathcal{F}} \left\{ \sup_{\nabla_t} \langle \hat{y} - f, \nabla_t \rangle + \left\langle f, -R_t - \sum_{s=1}^{t-1} \nabla_s \right\rangle \right\} \\
&= \underset{\hat{y}\in\mathcal{F}}{\text{argmin}} \sup_{f\in\mathcal{F}} \left\{ \|\hat{y} - f\|_\infty - \left\langle f, R_t + \sum_{s=1}^{t-1} \nabla_s \right\rangle \right\}
\end{aligned}$$

Example: Finite Experts

- Very similar to ℓ_1/ℓ_∞, think about subtracting -1 from every loss, makes no difference for regret
- But then ℓ_1/ℓ_∞ is same as finite experts
- Algorithm : Round t draw $R_t \sim N(0, (n-t)I_{|\mathcal{F}|})$

$$\hat{y}_t = \operatorname*{argmin}_{i \in [d]} \sum_{j=1}^{t} \ell(i, z_t) + R_t[i]$$

- Bound : $\mathbb{E}\left[\mathrm{Reg}_n\right] \le \frac{1}{n}\mathbf{Rel}_n(\cdot) = O\left(\sqrt{\frac{\log|\mathcal{F}|}{n}}\right)$

EXAMPLE: ONLINE SHORTEST PATH

- Graph $G = (V, E)$, source node S and destination node D.
- Every round, we need to pick a path from S to D
- Adversary picks a delay on every edge $W : E \mapsto [0, 1]$
- Learner suffers delay on path chosen which is sum of delays on edges of the path
- Experts bound $|E|\sqrt{\frac{|V| \log |V|}{n}}$
- However naive time complexity $O(\#\text{paths})$

Example: Online Shortest Path

- Can view it as a different online linear optimization problem
- $\mathcal{F} = \{f \in \{0,1\}^{|E|} : f \text{ is a path}\}$
- $\mathbf{D} = [0,1]^{|E|}$ the delays on each edge.
- Random playout condition satisfied by distribution $D = N(0,1)$
- Algorithm: Draw $R_t \sim N(0, (n-t)I_{|E|})$,

$$\text{path}_t = \underset{f \in \mathcal{F}}{\text{argmin}} \left\langle f, \sum_{j=1}^{t-1} \nabla_j + R_t \right\rangle$$

- That is solve shortest path algorithm with delay on edge $e \in E$ given by $\sum_{j=1}^{t-1} \nabla_j[e] + R_t[e]$
- Can be solves in poly-time using Bellman-ford algorithm.

Whirlwind Tour: Itinerary

1. Learning with partial feedback (bandits and more)
2. Learning and game theory
3. Learning and optimization
4. Algorithmic stability and learnability
5. Learning and Minimax theorem

Bandit Linear Optimization

For $t = 1$ to n

Learner picks move $\hat{\mathbf{y}}_t \in \mathcal{F}$

Adversary simultaneously picks ∇_t

Learner suffers loss $\langle \hat{\mathbf{y}}_t, \nabla_t \rangle$ and only gets to see this (not ∇_t)

End For

Bandits: main idea

Access to a full information algorithm :

- Obtain $\hat{\mathbf{y}}_t$ from the full information algorithm
- Randomize the move such that $\mathbb{E}[\hat{y}_t] = \mathbf{y}_t$ (or approximately)
- Play $\hat{y}_t$ and Receive $\langle \hat{y}_t, \nabla_t \rangle$
- Build unbiased estimate of ∇_t based on $\langle \hat{y}_t, \nabla_t \rangle$ $\quad (\mathbb{E}_{\hat{y}_t}\left[\tilde{\nabla}_t\right] = \nabla_t)$
- Feed $\tilde{\nabla}_t$ to the full information algorithm for round t

Getting Unbiased Estimates

We want $\mathbb{E}\left[\tilde{\nabla}_t\right] = \nabla_t$

- For multi-armed bandit : $\tilde{\nabla}_t = \frac{\langle \hat{y}_t, y_t \rangle}{q_t(\hat{y}_t)} \hat{y}_t$
- For $\mathcal{F}/\mathbf{D} = \ell_2^d/\ell_2^d$: pick at time t any orhtonormal basis

  $\hat{y}_t = \hat{\mathbf{y}}_t + \epsilon_t r_{i_t} e_{i_t}, \quad \tilde{\nabla}_t = \frac{\epsilon_t \langle \hat{y}_t, y_t \rangle}{r_{i_t} q_t(i_t)} e_{i_t}$

  r_{i_t} is specific to choice picked and ensures that $\hat{y}_t \in \mathcal{F}$

BANDITS: MAIN IDEA

$$\begin{aligned}
n\,\mathbb{E}\left[\mathrm{Reg}_n\right] &= \mathbb{E}\left[\sum_{t=1}^{n}\langle \hat{y}_t, \nabla_t\rangle\right] - \inf_{f\in\mathcal{F}}\sum_{t=1}^{n}\langle f, \nabla_t\rangle \\
&= \mathbb{E}\left[\sum_{t=1}^{n}\langle \hat{\mathbf{y}}_t, \nabla_t\rangle\right] - \inf_{f\in\mathcal{F}}\sum_{t=1}^{n}\langle f, \nabla_t\rangle + \mathrm{approx}_n \\
&= \mathbb{E}\left[\sum_{t=1}^{n}\left\langle \hat{\mathbf{y}}_t, \mathbb{E}\left[\tilde{\nabla}_t\right]\right\rangle\right] - \inf_{f\in\mathcal{F}}\sum_{t=1}^{n}\left\langle f, \mathbb{E}\left[\tilde{\nabla}_t\right]\right\rangle + \mathrm{approx}_n \\
&\leq \mathbb{E}\left[\underbrace{\sum_{t=1}^{n}\langle \hat{\mathbf{y}}_t, \tilde{\nabla}_t\rangle - \inf_{f\in\mathcal{F}}\sum_{t=1}^{n}\langle f, \tilde{\nabla}_t\rangle}_{\text{Regret bounded for full information algo}}\right] + \mathrm{approx}_n
\end{aligned}$$

BANDITS: MAIN IDEA

$$n\,\mathbb{E}\left[\mathrm{Reg}_n\right] = \eta\mathbb{E}\left[\sum_{t=1}^{n} \left\|\tilde{\nabla}_t\right\|_*^2\right] + \frac{1}{\eta}\Delta_R(\hat{y}_1|f^*) + \mathrm{approx}_n$$

- For crude bound: $\mathbb{E}[\hat{y}_t] \approx \hat{\mathbf{y}}_t$ got by mixing γ times the uniform distribution (for multiarmed bandit)
- $\mathrm{approx}_n = \gamma n, \quad \mathbb{E}\left[\sum_{t=1}^{n} \left\|\tilde{\nabla}_t\right\|_*^2\right] = \frac{d^2 n}{\gamma}$
- Gives sub-optimal $O(n^{-1/3})$ rate for regret

Bandits: main idea (refined)

$$n\,\mathbb{E}\left[\mathrm{Reg}_n\right] = \eta\mathbb{E}\left[\sum_{t=1}^{n}\left\|\tilde{\nabla}_t\right\|_t^2\right] + \frac{1}{\eta}\Delta_R(\hat{y}_1|f^*)$$

1. Instead of choosing any strongly convex R for MD choose one that blows up near the boundary of $\mathcal{F}$
   Eg. entropy, self-concordant barriers etc.
2. Refine MD proof to get local norms
3. Now can do exact $\mathbb{E}\left[\tilde{\nabla}_t\right] = \nabla_t$
4. $\mathbb{E}\left[\left\|\tilde{\nabla}_t\right\|_t^2\right] \leq \tilde{O}(d)$
5. Get $\tilde{O}\left(\sqrt{\frac{d}{n}}\right)$ rates

Other partial information problems

1. Contextual Bandit: Input variable x_t's and hypothesis class $\mathcal{F}$, only get loss on our prediction and not other options
2. Partial monitoring: Fixed loss matrix L and observation matrix H $(M \times N)$.
   - We pick action $i_t \in [M]$
   - Adversary picks action $j_t \in [N]$
   - We pay loss $L[i_t, j_t]$ but only observe $H[i_t, j_t]$
   - Rates either $n^{-1/2}$ or $n^{-1/3}$ or $O(1)$ based on conditions on H and L
3. Online zeroth order convex Lipchitz optimization:
   - Loss is convex and Lipschitz in out action (not just linear)
   - Key idea : Unbiased estimate of gradient

$$\nabla L_t(\hat{\mathbf{y}}_t) = \mathbb{E}\left[d \frac{\hat{\mathbf{y}}_t + \delta \mathbf{u}_t}{\delta} \mathbf{u}_t \right]$$

   - Open problem: Lower bound $O(n^{-1/2})$ upper bound $O(n^{-1/3})$

Learning and Games: Zero-sum two player games

- Payoff matrix L of size $M \times N$
- Player I plays a regret minimizing strategy, average pay off approximately as good as minimax value
- If both players play regret minimizing strategy, average moves converge to minimax equilibiria

Learning and Games: Zero-sum two player games

$$\frac{1}{n}\sum_{t=1}^{n} a_t^\top L b_t \le \min_{a \in \Delta_M} \frac{1}{n}\sum_{t=1}^{n} a^\top L b_t + \sqrt{\frac{\log M}{n}}$$

$$\le \sup_{b \in \Delta_N} \min_{a \in \Delta_M} a^\top L b + \sqrt{\frac{\log M}{n}}$$

$$\le \text{minimax value} + \sqrt{\frac{\log M}{n}}$$

Learning and Games: Multi-player games

- Correlated equilibrium: Joint distribution q^* on moves of players is said to be a correlated equilibrium if for every player $i \in [K]$, and any $\pi_i : [N_i] \mapsto [N_i]$

$$\mathbb{E}_{(a_1,\ldots,a_K)\sim q^*}\left[L_i(a_i, a^{-i})\right] \le \mathbb{E}_{(a_1,\ldots,a_K)\sim q^*}\left[L_i(\phi_i(a_i), a^{-i})\right]$$

- Minimize swap regret (mixed action) for each player:

$$\frac{1}{n}\sum_{t=1}^{n} L_i(a_t, a_t^{-i}) - \inf_{\phi} \frac{1}{n}\sum_{t=1}^{n} L_i(\phi(a_t), a_t^{-i}) \to 0$$

- The average moves of players converge to correlated equilibrium
- Generalize to any set of mappings Φ

Learning and Optimization: Convex optimization

- Regret minimization strategy applied to same function for the n steps provides an optimization algorithm
- Oracle efficiency
  1. Efficiency measured by counting number of local oracle calls to get desired sub-optimality level
  2. Eg. number of gradient calculations, function evaluations, higher derivatives etc.
- $\mathcal{F}$ is a d dimensional convex set: convex Lipchitz function
  - Efficiency bounded by $d \log(1/\epsilon)$
  - What about when d is large, say order $1/\epsilon^2$
  - MD and online methods optimal: proof through seq. Rademacher
- Statistical learning is stochastic convex optimization

Learning and Optimization: approximation algorithms

- Consider problem of form

$$\underset{f \in \mathcal{G}}{\operatorname{argmax}} \quad c^\top f$$
$$\forall i \in [d] \;\; G_i(f) \leq 1$$

  G_i's are convex and easy to optimize over $\mathcal{G}$ (Eg. linear)
- Say F^* the value of the above (not known, do binary search)
- Say $\mathcal{G}$ is an easy polytope (Eg. linear equalities, flow polytope)
- Define $\mathcal{F} = \{f \in \mathcal{G} : c^\top f = F^*\}$
- Solution can be seen as:

$$f^* = \underset{f \in \mathcal{F}}{\operatorname{argmin}} \max_{i \in [d]} G_i(f)$$

Learning and Optimization: approximation algorithms

- Run experts algorithm over d experts (for the d constraints) and returns $q_t \in \Delta_d$ a distribution over (negative) constraints
- On each round return $f_t = \mathrm{argmin}_{f \in \mathcal{F}} \mathbb{E}_{i_t \sim q_t} \left[G_{i_t}(f) \right]$ (assume this is easy to do)
- We claim $\bar{f}_n = \frac{1}{n} \sum_{t=1}^{n} f_t$ is a solution that only mildly violates constraints and obtains optimal value F^*.
- Now if we can find an $f_0 \in \mathcal{G}$ s.t. $c^\top > 0$ and $\forall i \in [d],\ G_i(f_0) \leq 1 - \gamma$
- Then for appropriate α, $\hat{f}_t = (1-\alpha)\bar{f}_n + \alpha f_0$ is a $(1-\epsilon)$ approximate solution and satisfies constraints

Regret bound implies:

$$\begin{aligned}
\max_{i\in[d]} \frac{1}{n}\sum_{t=1}^{n} G_i(f_t) &\leq \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}_{i_t\sim q_t}\left[G_{i_t}(f_t)\right] + \sqrt{\frac{\log d}{n}} \\
&= \frac{1}{n} \operatorname*{argmin}_{f_t\in\mathcal{F}} \sum_{t=1}^{n} \mathbb{E}_{i_t\sim q_t}\left[G_{i_t}(f_t)\right] + \sqrt{\frac{\log d}{n}} \\
&\leq \frac{1}{n}\sum_{t=1}^{n} \mathbb{E}_{i_t\sim q_t}\left[G_{i_t}(f^*)\right] + \sqrt{\frac{\log d}{n}} \\
&\leq 1 + \sqrt{\frac{\log d}{n}}
\end{aligned}$$

Learning and Optimization: approximation algorithms

Regret bound implies:

$$
\begin{aligned}
1 + \sqrt{\frac{\log d}{n}} &\geq \max_{i \in [d]} \frac{1}{n} \sum_{t=1}^{n} G_i(f_t) \\
&\geq \max_{i \in [d]} G_i \left(\frac{1}{n} \sum_{t=1}^{n} f_t \right)
\end{aligned}
$$

Choosing n the number of iterations to run to appropriately with ϵ gives us a solution that ϵ-violates the constraints

Learning and Optimization: approximation algorithms

- This approach was key to getting $1 - \epsilon$ approximate max-flow solution in time
$$\tilde{O}\left(|E|\,|V|^{1/3}\epsilon^{-11/3}\right)$$
- More careful and refined versions finally lead to $\tilde{O}(|E|/\epsilon^2)$ type rates.

Stability and Learnability

- Bounds for statistical (also for online) learning can be shown through algorithmic stability
- Uniform leave one out stability: for all sample S of size n and any $i \in [n]$

$$\left|\ell(A(S^{\setminus i}), z_i) - \ell(A(S), z_i)\right| \leq \epsilon_{stable}(n)$$

- Approximate ERM:

$$\frac{1}{n}\sum_{t=1}^{n}\ell(A(S), z_t) - \inf_{f \in \mathcal{F}} \frac{1}{n}\sum_{t=1}^{n}\ell(f, z_t) \leq \epsilon_{aerm}(n)$$

Stability and Learnability

- Stability implies that for the algorithm

$$\mathbb{E}_S\left[\left|\frac{1}{n}\sum_{t=1}^{n}\ell(A(S),z_t) - L_D(A(S))\right|\right] \to 0$$

- AERM + above implies

$$\mathbb{E}_S\left[L_D(A(S)) - \inf_{f\in\mathcal{F}}\frac{1}{n}\sum_{t=1}^{n}\ell(f,z_t)\right] \to 0$$

A general learning problem is learnable if and only if there exists a stable AERM

Online Learning and Minimax Theorem

- Von Neumann Minimax theorem:

  $\min_{a\in\Delta_M} \max_{b\in\Delta_N} a^\top L b = \max_{b\in\Delta_N} \min_{a\in\Delta_M} a^\top L b$

- Sion's Minimax theorem: $\mathcal{A}$ compact and convex, $\mathcal{B}$ convex, $\ell(a,\cdot)$ quasi-concave and upper semi-continuous, $\ell(\cdot,b)$ quasi-convex and lower semi-continuous

$$\inf_{a\in\mathcal{A}} \sup_{b\in\mathcal{B}} \ell(a,b) = \sup_{b\in\mathcal{B}} \inf_{a\in\mathcal{A}} \ell(a,b)$$

- For infinite dimensional spaces, unit ball not compact, but linear form minimax could still hold: $\mathcal{A}$ convex weakly compact subset of Banach space

$$\inf_{a\in\mathcal{A}} \sup_{b\in\mathcal{B}} \langle a,b\rangle = \sup_{b\in\mathcal{B}} \inf_{a\in\mathcal{A}} \langle a,b\rangle$$

Online Learning and Minimax Theorem

- Key to proving regret bounded by $\mathcal{R}_n^{sq}(\ell \circ \mathcal{F})$ was that minimax theorem holds:

$$\inf_{q \in \Delta(\mathcal{F})} \sup_{z \in \mathcal{Z}} \mathbb{E}_{f \sim q} [\ell(f,z)] = \inf_{p \in \Delta(\mathcal{Z})} \sup_{f \in \mathcal{F}} \mathbb{E}_{z \sim p} [\ell(f,z)]$$

- Is this a necessary condition?

$$
\begin{aligned}
\mathcal{V}_n^{sq} &= \inf_{\pi} \sup_{\tau} \mathbb{E}_{\pi} \left\{ \frac{1}{n} \sum_{t=1}^{n} \ell(f_t, z_t) - \inf_{f \in \mathcal{F}} \frac{1}{n} \sum_{t=1}^{n} \ell(f, z_t) \right\} \\
&= \inf_{\pi} \sup_{\tau} \mathbb{E}_{\pi} \left\{ \frac{1}{n} \sum_{t=1}^{n} \ell(f_t, z_t) - \inf_{f \in \mathcal{F}} \mathbb{E}_{z \sim \hat{p}} \ell(f, z) \right\} \\
&\geq \inf_{\pi} \sup_{\tau} \mathbb{E}_{\pi} \left\{ \frac{1}{n} \sum_{t=1}^{n} \ell(f_t, z_t) - \sup_{p} \inf_{f \in \mathcal{F}} \mathbb{E}_{z \sim p} \ell(f, z) \right\} \\
&= \left\{ \inf_{\pi} \sup_{\tau} \mathbb{E}_{\pi} \frac{1}{n} \sum_{t=1}^{n} \ell(f_t, z_t) \right\} - \left\{ \sup_{p} \inf_{f \in \mathcal{F}} \mathbb{E}_{z \sim p} \ell(f, z) \right\} \\
&= \inf_{q \in \Delta(\mathcal{F})} \sup_{\mathbf{z} \in \mathcal{Z}} \mathbb{E}_{f \sim q} \ell(f, z) - \sup_{p \in \Delta(\mathcal{Z})} \inf_{f \in \mathcal{F}} \mathbb{E}_{z \sim p} \ell(f, z)
\end{aligned}
$$

Online Learning and Minimax Theorem

If minimax theorem does not hold, then $\exists c > 0$ s.t.

$$\inf_{q \in \Delta(\mathcal{F})} \sup_{z \in \mathcal{Z}} \mathbb{E}_{f \sim q} \ell(f, z) \geq \sup_{p \in \Delta(\mathcal{Z})} \inf_{f \in \mathcal{F}} \mathbb{E}_{z \sim p} \ell(f, z) + c$$

Hence,

$$\mathcal{V}_n(\mathcal{F}) \geq c > 0$$