Regret Bounds of a Distributed Saddle Point Algorithm

Alec Koppel, Felicia Jakubiec, Alejandro Ribeiro

Dept. of Electrical and Systems Engineering, University of Pennsylvania

{akoppel, life, aribeiro}@seas.upenn.edu

Int. Conf. on Acoustics, Speech, and Signal Processing

Brisbane, AU, April 24, 2015

Online Learning

- Online $\Rightarrow$ information received sequentially
- Framework for describing family of adaptive algorithms
- Classical offline model fitting strategy:

  $\Rightarrow$ minimize residual error over data set all at once
- Online learning inverts this procedure:

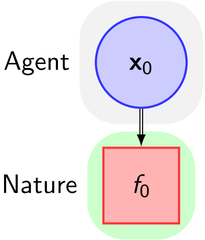
  $\Rightarrow$ repeatedly adjust model based on new information

- Regret $\Rightarrow$ performance metric for online algorithms

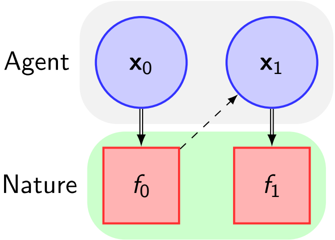
  $\Rightarrow$ Measures no. of mistakes against a fixed optimal offline learner

  $\Rightarrow$ How well is learner adapting to new information?

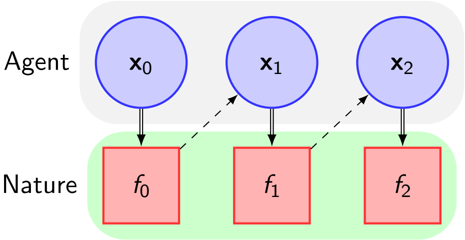
Online Learning



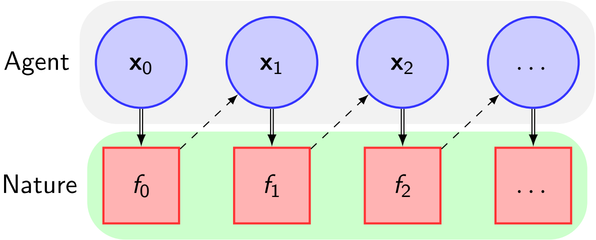
Online Learning



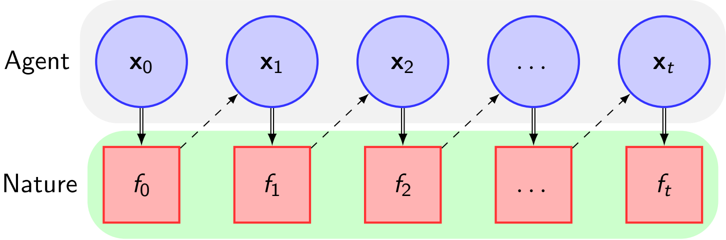
Online Learning



Online Learning



Online Learning



- Repeated game over a convex $X \subset \mathbb{R}^p$
- At the t_{th} round, agent plays $\mathbf{x}_t \in X$, Nature reveals $f_t : X \to \mathbb{R}$

  $\Rightarrow$ Suffer arbitrary independent (antagonistic) convex loss $f_t(\mathbf{x}_t)$

Regret

- Regret $\Rightarrow$ performance metric for online learning

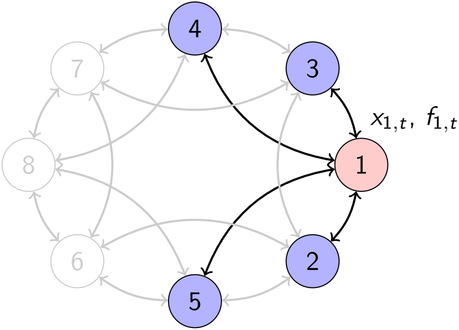
$$\mathbf{Reg}_T := \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x}^*)$$

- For fixed T, $x^* = \text{argmin}_{x \in X} \sum_{t=1}^{T} f_t(\mathbf{x})$ is *offline* learner
  - $\Rightarrow$ Price for causal operation
  - $\Rightarrow$ How much we pay for not being clairvoyant

- Goal: $\mathbf{Reg}_T / T \to 0$ as $T \uparrow$, online gradient descent (Zinkevich, '03)

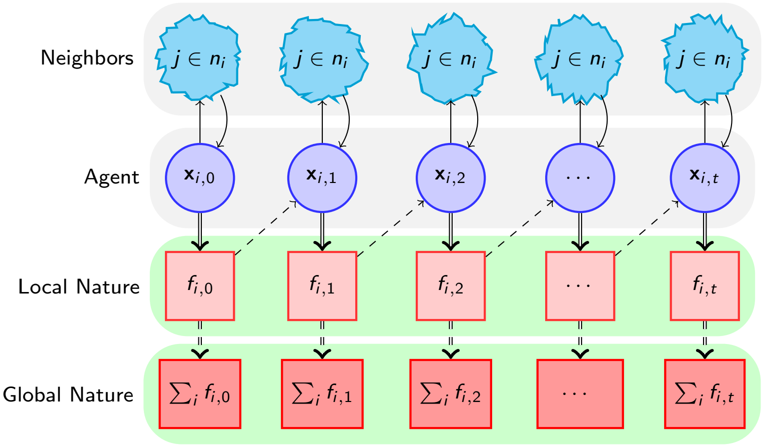
Repeated Network Games

- Network $\mathcal{G} = (V, \mathcal{E})$
  $\Rightarrow |V| = N$, $|\mathcal{E}| = M$
- Neighborhood of agent i
  $\Rightarrow n_i = \{j : (j, i) \in \mathcal{E}\}$
- Repeated game at agent i, time t
  $\Rightarrow$ action $\tilde{\mathbf{x}}_{i,t}$ $\Rightarrow$ local loss $f_{i,t}$



- Minimize only local loss
  $\Rightarrow$ decoupled local learning
- Instead, each agent i aims to
  $\Rightarrow$ minimize global loss $f_t(\tilde{\mathbf{x}}_t) = \sum_{i=1}^{N} f_{i,t}(\tilde{\mathbf{x}}_t)$
  $\Rightarrow$ only observes local loss $f_{i,t}$ $\Rightarrow$ collaborate with other agents

Local information, Global Loss



- Each node observes local loss and neighbors' information
  $\Rightarrow$ Minimize network-wide loss $\Rightarrow$ learn global info
- Distribution-free: Nature arbitrarily different at distinct nodes

Learning in Networks is Possible

- Learning in networks $\Rightarrow$ regret comparable to centralized case
- Prior approaches: consensus protocol (Trianos 2012, Yan 2011)
  $\Rightarrow$ Inspired by distributed control $\Rightarrow$ swarm and synchronization
- Claim: New networked online learning strategy
  $\Rightarrow$ saddle point alg.
- Rather than weighted averaging, penalize agent disagreement
  $\Rightarrow$ primal-dual algorithm $\Rightarrow$ seek saddle point of online Lagrangian
- Proposed method achieves networked online learning goal:
  $\Rightarrow$ $\mathbf{Reg}_T/T \to 0$ as $T \uparrow$ at rate $O(1/\sqrt{T})$

Formulating Repeated Network Games

- Coordinated Regret: $\Rightarrow$ all agents play common $\tilde{\mathbf{x}}_t$

$$\mathbf{Reg}_T^C = \sum_{t=1}^{T} \sum_{i=1}^{N} f_{i,t}(\tilde{\mathbf{x}}_t) - \sum_{t=1}^{T} \sum_{i=1}^{N} f_{i,t}(\mathbf{x}^*).$$

- Distributed games $\Rightarrow$ agents autonomous: $\tilde{\mathbf{x}}_{i,t} \neq \tilde{\mathbf{x}}_{j,t}, j \neq i$
- Uncoordinated Regret

$$\mathbf{Reg}_T^U = \sum_{t=1}^{T} \sum_{i=1}^{N} f_{i,t}(\tilde{\mathbf{x}}_{i,t}) - \sum_{t=1}^{T} \sum_{i=1}^{N} f_{i,t}(\mathbf{x}^*).$$

- Convexity implies unique global offline $\mathbf{x}^*$ $\Rightarrow$ nodes should agree
- $\mathbf{Reg}_T^U$ fails to incentivize collaboration

Networked Regret

- **Local Regret** of node i of distributed online algorithm

$$\mathbf{Reg}_T^i = \sum_{t=1}^{T} \sum_{j=1}^{N} f_{j,t}(\tilde{\mathbf{x}}_{i,t}) - \sum_{t=1}^{T} \sum_{j=1}^{N} f_{j,t}(\mathbf{x}^*).$$

$\Rightarrow$ $\mathbf{x}^* = \operatorname{argmin}_{\mathbf{x}} \sum_{t=1}^{T} \sum_{j=1}^{N} f_{j,t}(\mathbf{x})$ is the global batch learner
$\Rightarrow$ Quality of node i's prediction at others' losses

- **Global Networked Regret**

$$\mathbf{Reg}_T := \frac{1}{N} \sum_{i=1}^{N} \mathbf{Reg}_T^i = \frac{1}{N} \sum_{t=1}^{T} \sum_{i,j=1}^{N} f_{j,t}(\tilde{\mathbf{x}}_{i,t}) - \sum_{t=1}^{T} \sum_{j=1}^{N} f_{j,t}(\mathbf{x}^*),$$

- Networked online learning goal: $\mathbf{Reg}_T^i / T, \mathbf{Reg}_T / T \to 0$ as $T \uparrow$
$\Rightarrow$ Measures how well agents learn global information

Agent Agreement and Lagrangian Relaxation

- Convexity of $f_{i,t} \implies$ unique globally optimal offline strategy $\mathbf{x}^*$
  - $\Rightarrow$ Node predictions should coincide at optimality.
- At each t we want to enforce $\mathbf{x}_{i,t} = \mathbf{x}_{j,t}$ for $j \in n_i$, or $\mathbf{C}\mathbf{x}_t = \mathbf{0}$.
  - $\Rightarrow$ $\mathbf{C}$ is node-edge incidence matrix of $\mathcal{G}$.
  - $\Rightarrow$ $\mathbf{x}_t$ is stacked version of $\mathbf{x}_{i,t}$.
- Constraint enforcement requires global coordination
  - $\Rightarrow$ Lagrangian relaxation allows distributed computation
- Online Lagrangian for networked learning problem:

$$\mathcal{O}_t(\mathbf{x}_t, \boldsymbol{\lambda}_t) = \sum_{i=1}^{N} f_{i,t}(\mathbf{x}_{i,t}) + \boldsymbol{\lambda}_t^\intercal \mathbf{C}\mathbf{x}_t$$

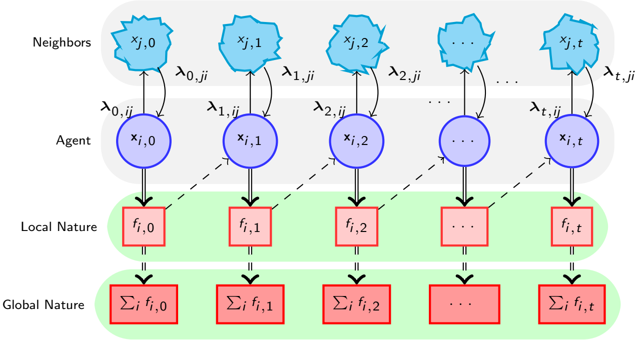
- Convex/concave function in the primal/dual variables, respectively

Arrow-Hurwicz Saddle Point Method

- Arrow-Hurwicz saddle point method on online Lagrangian $\mathcal{O}_t(\mathbf{x}_t, \boldsymbol{\lambda}_t)$
  - $\Rightarrow$ Primal Lagrangian subgradient descent
  - $\Rightarrow$ Dual Lagrangian subgradient ascent
- Algorithm formulation:

$$\mathbf{x}_{t+1} = \mathcal{P}_X[\mathbf{x}_t - \epsilon \nabla_{\mathbf{x}} \mathcal{O}_t(\mathbf{x}_t, \boldsymbol{\lambda}_t)]$$
$$\boldsymbol{\lambda}_{t+1} = \mathcal{P}_\Lambda[\boldsymbol{\lambda}_t + \epsilon \nabla_{\boldsymbol{\lambda}} \mathcal{O}_t(\mathbf{x}_t, \boldsymbol{\lambda}_t)]$$

- Initialize $\boldsymbol{\lambda}_1 = \mathbf{0}$ for $\boldsymbol{\lambda}_t$ to remain in the image of $\mathbf{C}$
  - $\Rightarrow$ Required for bounded dual subgradients and constraint slacks



- Primal step at each node i $\Rightarrow$ gradient with local penalty

$$\mathbf{x}_{i,t+1} = \mathcal{P}_X\left[\mathbf{x}_{i,t} - \epsilon\left(\nabla_{\mathbf{x}_i} f_{i,t}(\mathbf{x}_{i,t}) + \sum_{j\in n_i} \lambda_{ij,t} - \lambda_{ji,t}\right)\right]$$

- Dual step at each edge (i,j) $\Rightarrow$ increase price of disagreement

$$\boldsymbol{\lambda}_{ij,t+1} = \mathcal{P}_{\Lambda_{ij}}\left[\boldsymbol{\lambda}_{ij,t} + \epsilon\left(\mathbf{x}_{i,t} - \mathbf{x}_{j,t}\right)\right]$$

Uncoordinated Regret Bound

Lemma

The saddle point algorithm with dual initialization $\boldsymbol{\lambda}_1 = \mathbf{0}$ and constant step size $\epsilon = 1/\sqrt{T}$, achieves uncoordinated regret bound

$$\mathbf{Reg}_T^U + \sum_{t=1}^{T} \boldsymbol{\lambda}^T \mathbf{C}\mathbf{x}_t \leq \frac{\sqrt{T}}{2} \left(\|\mathbf{x}_1 - \mathbf{x}^*\|^2 + \|\boldsymbol{\lambda}\|^2 + L_{\mathbf{x}}^2 + L_{\boldsymbol{\lambda}}^2 \right).$$

for all $\boldsymbol{\lambda} \in \Lambda$.

- $\mathbf{x}^*$ the global batch learner,
- Agents can bound uncoordinated regret without collaboration
- Bound in lemma includes disagreement penalty $\sum_{t=1}^{T} \boldsymbol{\lambda}^T \mathbf{C}\mathbf{x}_t$.

$\Rightarrow$ Relate uncoordinated and networked regret via dual variable.

Global Networked Regret Bound

Theorem

The saddle point alg. sequence with initialization $\boldsymbol{\lambda}_1 = \mathbf{0}$ *and constant step size* $\epsilon = 1/\sqrt{T}$ *achieves the Global networked regret bound*

$$\mathbf{Reg}_T \;\le\; \frac{\sqrt{T}}{2}\left(\|\mathbf{x}_1 - \mathbf{x}^*\|^2 + MC_{\boldsymbol{\lambda}}^2 + L_{\mathbf{x}}^2 + L_{\boldsymbol{\lambda}}^2\right) \;=\; O(\sqrt{T}).$$

- $\mathbf{Reg}_T/T \to 0$ with $T \uparrow$, learning constant depends on . . .
  $\Rightarrow$ Network size N and diameter D
  $\Rightarrow$ Initialization, Lipschitz constant K, gradient bounds $L_{\mathbf{x}}$, $L_{\boldsymbol{\lambda}}$
- $C_{\boldsymbol{\lambda}}$ must satisfy $C_{\boldsymbol{\lambda}} \ge DNK + 1$ $\Rightarrow$ dual set projection
- Comparable to centralized regret bound of online gradient descent

Local Regret Bound

Theorem

If we run the saddle point alg. for T rounds with initialization $\boldsymbol{\lambda}_1 = \mathbf{0}$ with constant step size $\epsilon = 1/\sqrt{T}$

Local regret bound of node i:

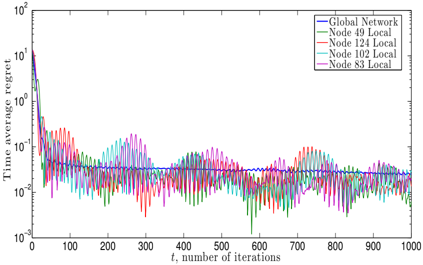
$$\mathbf{Reg}_T^i \;\leq\; \frac{\sqrt{T}}{2}\left(\|\mathbf{x}_1 - \mathbf{x}^*\|^2 + MC_{\boldsymbol{\lambda}}^2 + L_{\mathbf{x}}^2 + L_{\boldsymbol{\lambda}}^2\right) \;=\; O(\sqrt{T}).$$

- $\mathbf{x}^*$ the global batch learner,
- Each agent learns global info $\Rightarrow$ same rate as network in aggregate
- Learning constants $\Rightarrow$ initialization, loss functions, and network

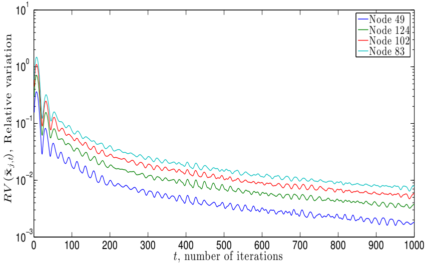
- $N = 200 \;\Rightarrow$ random subset
  - $\Rightarrow$ connected w.p. 0.2
- $\mathbf{H}_i \in \{1/p, 2/p, \ldots, 1\}$
  - $\Rightarrow$ Chosen uniformly randomly
- $T = 1 \times 10^3$, $\epsilon = 1/\sqrt{T} \approx 0.03$
- Signal/obs dims. $p = 10, q = 1$
- $\mathbf{w}_{i,t} \sim \mathcal{N}(0, \sigma^2 = 0.1), \mathbf{x} = \mathbf{1}$
- $\mathbf{Reg}_T/T, \mathbf{Reg}^i_T/T \to 0$

$$\mathrm{RV}(\bar{\mathbf{x}}_{j,t}) := \sum_{k=1}^{N} \|\bar{\mathbf{x}}_{j,t} - \bar{\mathbf{x}}_{k,t}\| / N \|\mathbf{x}^*\|.$$

- Individuals reach consensus.



(a) $\mathbf{Reg}_t/t$, vs. local regrets $\mathbf{Reg}^i_t/t$



(b) $\mathrm{RV}(\bar{\mathbf{x}}_{i,t})$ over iteration t

Conclusion

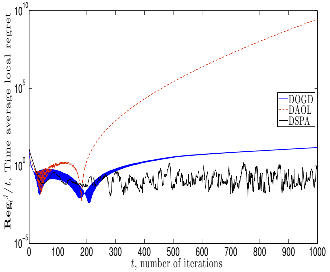
- Online learning $\Rightarrow$ networked settings
  $\Rightarrow$ Distribution free, Nature arbitrarily antagonistic
- Networked online learning as local regret minimization

- Derivation of saddle point algorithm
  $\Rightarrow$ Formulated online Lagrangian
  $\Rightarrow$ Arrow-Hurwicz saddle point method
- Theorem $\Rightarrow$ networked online learning guaranteed

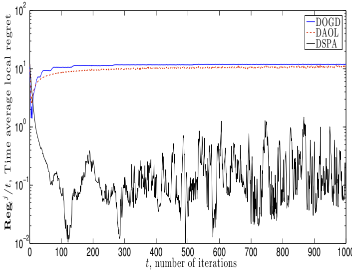
- Numerical analysis on distributed RLS
  $\Rightarrow$ Empirically achieve learning in networks

Algorithm Comparison

- Scalar obs. $\mathbf{y}_{i,t}$ $(q = 1)$, signal $\mathbf{x}$ dim. $p = 10$, arbitrary $j \in V$



(a) $\mathbf{Reg}_t^j/t$, local regret vs. t, $N = 50$ cycle



(b) $\mathbf{Reg}_t^j/t$, local regret vs. t, $N = 50$ grid

- Consensus methods
  - $\Rightarrow$ DAOL diverges
  - $\Rightarrow$ DOGD constant regret
- Consensus methods
  - $\Rightarrow$ constant networked regret
  - $\Rightarrow$ network agreement
  - $\not\Rightarrow$ global batch learner
- Saddle point method learns, more effective when . . .
  - $\Rightarrow$ each $f_{i,t}$ has unique minima $\Rightarrow$ signal avg. $\not\Rightarrow$ loss avg.