Online Learning with Predictable Sequences

Sasha Rakhlin Karthik Sridharan

University of Pennsylvania

June 12, 2013

Online Learning

Protocol:

For $t = 1, \ldots, T$
  Learner chooses $f_t \in \mathcal{F}$
  Nature chooses $x_t \in \mathcal{X}$

External regret:

$$\sum_{t=1}^{T} \ell(f_t, x_t) - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \ell(f, x_t)$$

How do we add prior knowledge about sequences?

Model:

sequence = predictable process + adversarial noise

$$x_t = M_t(x_1, \ldots, x_{t-1}) + \delta_t$$

(R., Sridharan, Tewari '11): Minimax analysis for constraints

$$\|x_t - M_t(x_1, \ldots, x_{t-1})\| \leq \sigma_t$$

Model:

sequence = predictable process + adversarial noise

$$x_t = M_t(x_1, \ldots, x_{t-1}) + \delta_t$$

(R., Sridharan, Tewari '11): Minimax analysis for constraints

$$\|x_t - M_t(x_1, \ldots, x_{t-1})\| \leq \sigma_t$$

Surprising: knowledge of σ_t's is not required:

$$\sum_{t=1}^{T} \langle f_t, x_t \rangle - \inf_{f \in \mathcal{F}} \sum_{t=1}^{T} \langle f, x_t \rangle \leq c \sqrt{\sum_{t=1}^{T} \|x_t - M_t(x_1, \ldots, x_{t-1})\|_*^2}$$

for the problem of online linear optimization $\ell(f, x) = \langle f, x \rangle$.

Examples

- Path length:

$$M_t(x_1, \ldots, x_{t-1}) = x_{t-1}$$

- Variance:

$$M_t(x_1, \ldots, x_{t-1}) = \frac{1}{t-1} \sum_{s=1}^{t-1} x_s$$

- Fading memory:

$$M_t(x_1, \ldots, x_{t-1}) = \sum_{s=1}^{t-1} \alpha_s x_s, \qquad \sum_{s=1}^{t-1} \alpha_s = 1, \qquad \alpha_s \geq 0$$

- Phases:

$$M_t(x_1, \ldots, x_{t-1}) = x_{t-k}$$

- More general setting:

$$M_t(I_1, \ldots, I_{t-1}, f_1, \ldots, f_{t-1}, q_1, \ldots, q_{t-1})$$

Full-Information Algorithms

Optimistic Follow the Regularized Leader

Input: $\mathcal{R}$ self-concordant barrier, learning rate $\eta > 0$.

At $t = 1, \ldots, T$, predict f_t, observe x_t, and update

$$f_{t+1} = \arg\min_{f \in \mathcal{F}} \eta \left\langle f, \sum_{s=1}^{t} x_s + M_{t+1} \right\rangle + \mathcal{R}(f)$$

Optimistic Mirror Descent Algorithm

Input: $\mathcal{R}$ 1-strongly convex w.r.t. $\|\cdot\|$, learning rate $\eta > 0$

At $t = 1, \ldots, T$, predict f_t and update

$$g_{t+1} = \operatorname{argmin}_{g \in \mathcal{F}} \eta \langle g, x_t \rangle + D_{\mathcal{R}}(g, g_t)$$

$$f_{t+1} = \operatorname{argmin}_{f \in \mathcal{F}} \eta \langle f, M_{t+1} \rangle + D_{\mathcal{R}}(f, g_{t+1})$$

For $M_t = x_{t-1}$ this is the method of (Chiang, Yang, Lee, Mahdavi, Lu, Jin, Zhu, 2012)

Learning the Process

Model selection: given a family $(M_t^\pi)_{t\geq 1}$, for $\pi \in \Pi$, can we obtain a regret bound of order

$$\inf_{\pi \in \Pi} \sqrt{\sum_{t=1}^{T} \|x_t - M_t^\pi\|_*^2}$$

Partial Information

Only observe $\langle f_t, x_t \rangle$ and only $M_t^{\pi_t}$ corresponding to the predictable process of $\pi_t \in \Pi$ we select at time t.

Proofs crucially rely on a multi-armed bandit algorithm with a regret bound in terms of *loss of best arm* (we could not find it in the literature).

Further Directions

- Predictable Sequences beyond online convex optimization
- Faster rates for problems where regret black box is used
- Relation to structured optimization and Mirror Prox (in submission)
- Coupling with regret against strategies (in progress)
- Follow-the-Perturbed Leader for predictable sequences (see arXiv)