Competitive Mirror Descent

Competitive mirror descent
Florian Schäfer, Anima Anandkumar, and Houman Owhadi
arXiv preprint https://arxiv.org/abs/2006.10179



Caltech

The Problem:

- Constrained Competitive Optimization:

$$\min_{x \in \mathcal{C}} f(x, y) \quad \min_{y \in \mathcal{K}} g(x, y)$$

- Convex problems: Find Nash equilibrium

$$(\bar{x}, \bar{y}) \in \mathcal{C} \times \mathcal{K}: \quad \begin{aligned} \bar{x} &= \operatorname{argmin}_{x \in \mathcal{C}} \ f(x, \bar{y}), \\ \bar{y} &= \operatorname{argmin}_{y \in \mathcal{K}} \ g(\bar{x}, y) \end{aligned}$$

- Non-convex problems: More complicated …

Caltech

Example: Lagrangian Duality

- Rewrite constraint as maximum:

$$\min_{\substack{x \in \mathbb{R}^m \\ h(x) \leq 0}} f(x) \Leftrightarrow \min_{x \in \mathbb{R}^m} \max_{\lambda \geq 0} f(x) + \lambda h(x)$$

- Replace by competitive problem

$$\min_{x \in \mathbb{R}^m} f(x) + \lambda h(x), \qquad \min_{\lambda \geq 0} \lambda h(x)$$

- Generalization to general conic constraints

Caltech

Example: Adversarial Learning

- In GAN: $\mathcal{G}$ and $\mathcal{D}$ compete

$$\min_{\mathcal{G}} f(\mathcal{G}, \mathcal{D}),\ \min_{\mathcal{D}} -f(\mathcal{G}, \mathcal{D})$$

Produced by StyleGan2 on
https://thispersondoesnotexist.com/

- Originally, loss f defined as

$$\mathbb{E}_{x\sim P_{\text{data}}}\left[\log\left(\mathcal{D}(x)\right)\right] + \mathbb{E}_{z\sim\mathcal{N}}\left[\log\left(1\ - \mathcal{D}\left(\mathcal{G}(z)\right)\right)\right]$$

Caltech

Constrained Competitive Optimization

	Single agent	Competitive
Unconstrained	Gradient Descent	???
Constrained	???	???

Constrained Competitive Optimization

	Single agent	Competitive
Unconstrained	Gradient Descent	**Competitive gradient descent**
Constrained	???	???

Competitive Gradient Descent

- Unconstrained competitive optimization

$$\min_x f(x,y), \qquad \min_y g(x,y)$$

- Gradient descent solves linear approximation

$$x_k - \eta \nabla f(x_k) = \operatorname{argmin}_x \, [Df(x_k)](x - x_k) + \frac{||x - x_k||^2}{2\eta}$$

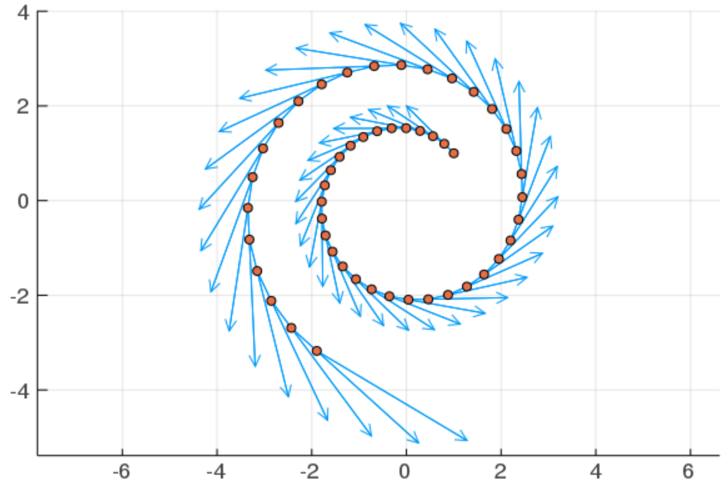
Caltech

Competitive Gradient Descent

- Naïve: Simultaneous Gradient Descent

$$x_{k+1} = x_k - \eta \nabla_x f(x_k, y_k)$$
$$y_{k+1} = y_k - \eta \nabla_y g(x_k, y_k)$$

- Fails even on bilinear
  $f(x, y) = xy = -g(x, y)$
  ``*Rock Paper Scissor…*"



Competitive Gradient Descent

- Idea: Nash equilibrium of bilinear approximation:

[Schaefer and Anandkumar, 2019]

$$x_{k+1} = \min_{x} \; [D_x f](x - x_k) + \left(x^\top - x_k^\top\right)[D_{xy}^2 f](y - y_k) + \left[D_y f\right](y - y_k) + \frac{||x - x_k||^2}{2\eta}$$

$$y_{k+1} = \min_{y} \; [D_x g](x - x_k) + \left(x^\top - x_k^\top\right)\left[D_{xy}^2 g\right](y - y_k) + \left[D_y g\right](y - y_k) + \frac{||y - y_k||^2}{2\eta}$$

- Has closed form solution

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - \eta \begin{pmatrix} \mathrm{Id} & \eta D_{xy}^2 f \\ \eta D_{yx}^2 g & \mathrm{Id} \end{pmatrix}^{-1} \begin{pmatrix} \nabla_x f \\ \nabla_y g \end{pmatrix}$$

Constrained Competitive Optimization

	Single agent	Competitive
Unconstrained	Gradient Descent	Competitive Gradient Descent
Constrained	Mirror Descent	Competitive Mirror Descent