

Fast Projection Onto Convex Smooth Constraints

Ilnura Usmanova

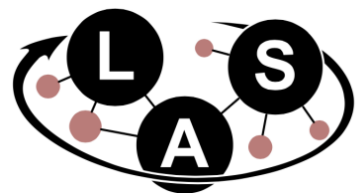
Maryam Kamgarpour

Andreas Krause

Kfir Yehuda Levy



International Conference on Machine Learning
2021



Learning &
Adaptive Systems

ETH zürich



Euclidean Projection Problem:

$$\begin{array}{ll} \min_{x \in \mathbb{R}^n} & \|x_0 - x\|^2 \\ \text{s.t.} & x \in \mathcal{K} \end{array}$$

$$\mathcal{K} := \{x \in \mathbb{R}^n : h_i(x) \leq 0, \forall i = 1, \dots, m\}$$

$h_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are smooth and convex

<u>Set</u>	<u>Approach</u>	<u>Computational complexity</u>	<u>Required oracle</u>
\mathcal{K} has a well defined self-concordant barrier	Interior point method (primal-dual IPM)	$O\left(n^3 \log \frac{n}{\varepsilon}\right)$	Second-order (Hessian)
$\mathcal{K} = \bigcap_{i=1}^m \mathcal{K}_i$	Alternating Directions Method of Multipliers (consensus ADMM)	$O\left(\frac{1}{\varepsilon}\right)$	Projection oracles for \mathcal{K}_i separately

Can we proceed faster in the high dimensional case? $O\left(nm^{2.5} \log^2(m/\varepsilon) + m^{3.5} \log(m/\varepsilon)\right)$

Warm-up:

Projection onto a single smooth constraint

Projection problem:

$$\begin{array}{ll} \min & \|x_0 - x\|^2 \\ \text{s.t.} & h(x) \leq 0 \end{array}$$

The Lagrangian:

$$L(x, \lambda) = \|x_0 - x\|^2 + \lambda h(x)$$

Strong duality:

$$\min_{x \in \mathbb{R}^d} \max_{\lambda \geq 0} L(x, \lambda) = \max_{\lambda \geq 0} \min_{x \in \mathbb{R}^d} L(x, \lambda)$$

Warm-up:

Projection onto a single smooth constraint

Dual problem: $\max_{\lambda \geq 0} d(\lambda)$

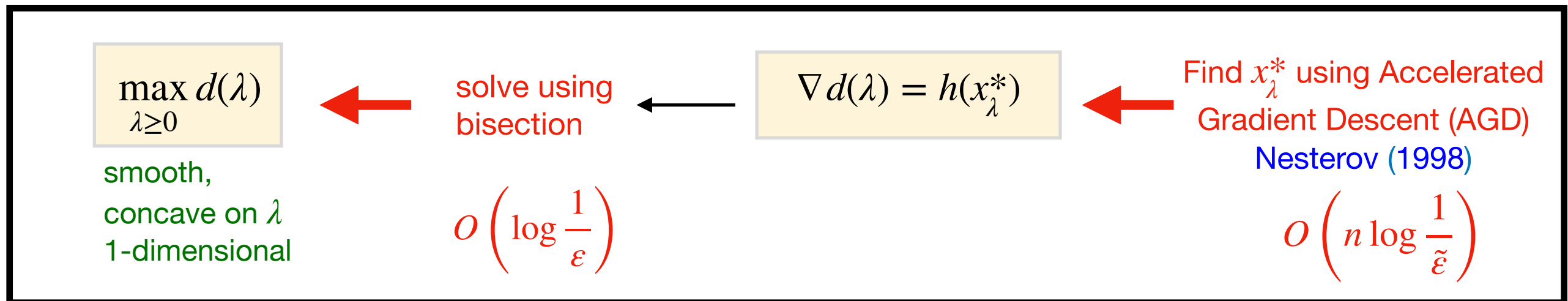
Dual function: $d(\lambda) = \min_{x \in \mathbb{R}^d} (\|x_0 - x\|^2 + \lambda h(x))$

Smooth & 2-strongly-convex

Due to strong convexity,
 x_λ^* is unique

$$d(\lambda) = \|x_0 - x_\lambda^*\|^2 + \lambda h(x_\lambda^*)$$

Gradient of the dual: $\nabla d(\lambda) = h(x_\lambda^*)$



Projection onto m smooth constraints

Projection problem:

$$\begin{aligned} \min & \|x_0 - x\|^2 \\ \text{s.t. } & \mathbf{h}(x) \leq 0 \\ & \mathbf{h}(x)_i = h_i(x), i = 1, \dots, m \end{aligned}$$

Dual problem:

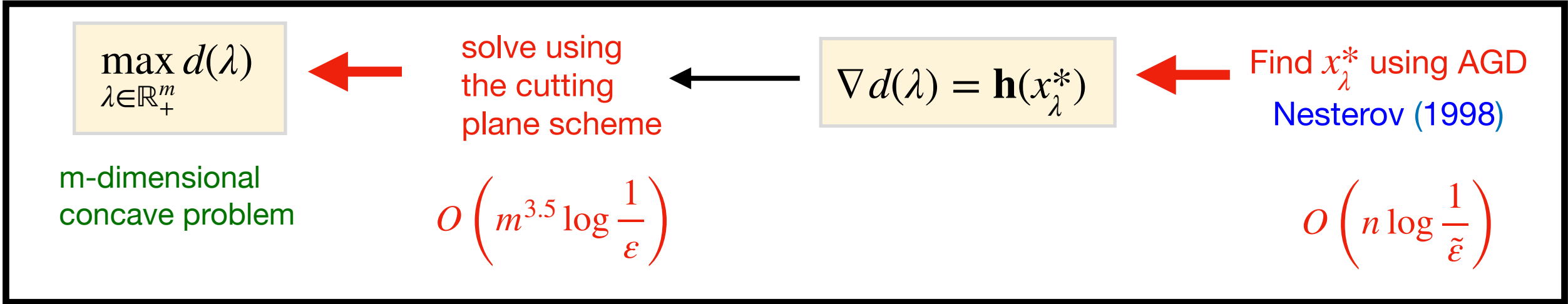
$$\max_{\lambda \in \mathbb{R}_+^m} d(\lambda)$$

$$d(\lambda) = \min_{x \in \mathbb{R}^d} \|x_0 - x\|^2 + \lambda^T \mathbf{h}(x)$$

Gradient of the dual objective:

$$\nabla d(\lambda) = \mathbf{h}(x_\lambda^*)$$

Fast Projection Method



Welcome to reading our paper!