Адаптивные методы семплирования по важности

Василий Новицкий, студент 4 курса ФУПМ
Научный руководитель: к.ф.-м.н. Ю.В. Максимов

МФТИ, ВЦ РАН, кафедра ИАД

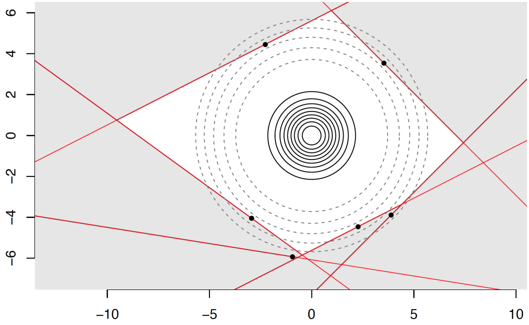
28 июня 2018

Постановка задачи

Случайный вектор $x \in \mathbb{R}^d$, $x \sim \mathcal{N}(\eta, \Sigma)$
Множество ограничений $H_j = \left\{x \mid w_j^\mathsf{T} x \geq \tau_j\right\}$

$$\mu = \mathrm{P}\left(x \in H\right), \; H = \cup_{j=1}^{J} H_j$$

Задача – оценить μ как можно быстрее



Метод Монте-Карло

Задача оценить μ, если даны $f(x) = \mathbb{I}(x \in H)$ и $p(x)$:

$$\mu = \mathrm{P}(x \in H) = \int f(x)p(x)dx$$

Оценка Монте-Карло:

$$\widehat{\mu} = \frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

$$\mathrm{Var}\ \widehat{\mu} = \frac{\sigma^2}{n},\ \sigma = \sqrt{\mu(1-\mu)}$$

$$\mathrm{cv} = \frac{\sigma/\sqrt{n}}{\mu} = \frac{\sqrt{\mu(1-\mu)}}{\mu\sqrt{n}} \sim \frac{1}{\sqrt{\mu n}}$$

$$\mathrm{cv} = 0.1,\ \mu \sim 10^{-5} - 10^{-7} \Rightarrow n = 100\mu^{-1} \sim 10^{7} - 10^{9}$$

Семплирование по важности

Идея семплирования по важности:

$$\mu = \int \frac{f(x)p(x)}{q(x)} q(x) dx$$

Оценка методом семплирования по важности:

$$\widehat{\mu}_q = \frac{1}{n} \sum_{i=1}^{n} \frac{f(x_i)p(x_i)}{q(x_i)}, \ x_i \sim q$$

Дисперсия

$$\mathrm{Var}\left(\widehat{\mu}_q\right) = \frac{1}{n}\left(\int \frac{(fp)^2}{q} dx - \mu^2\right) = \frac{1}{n} \int \frac{(fp - \mu q)^2}{q} dx$$

ALOE

Смесь условных распределений:

$$q_\alpha(x) = \sum_{j=1}^{J} \alpha_j q_j(x), \ \sum_{j=1}^{J} \alpha_j = 1$$

$$q_j(x) = \mathrm{P}\left(x \mid H_j\right) = \frac{p(x)H_j(x)}{P_j}$$

$$H_j(x) = \mathbb{I}(x \in H_j),\ H_{1..J}(x) = \mathbb{I}(x \in H), P_j = \mathrm{P}(H_j)$$

Веса: $\alpha_j^{ALOE} = P_j \left(\sum_{k=1}^{J} P_k \right)^{-1}$

Теорема (Owen, Maximov, Chertkov, 2017):

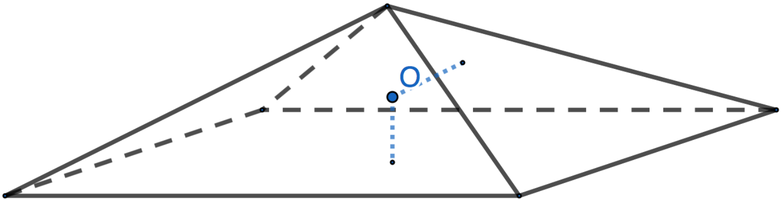
$$\mathrm{Var}\left(\widehat{\mu}_{ALOE}\right) \leq \frac{\mu^2}{n} \frac{J + J^{-1} - 2}{4}$$

Схема эксперимента

$$J = 1000 \text{ – число граней}$$

$$P_j = const \approx 3 \times 10^{-5}$$

$$\theta = 0.01 \text{ – угол при основании пирамиды}$$



Адаптивные веса

$$\mathrm{Var}\left(\widehat{\mu}_\alpha\right) = \frac{1}{n}\left(\int \frac{H_{1..J}(x)p(x)}{\sum_{j=1}^{J}\alpha_j H_j(x) P_j^{-1}}dx - \mu^2\right) \to \min$$

$$\sum_{j=1}^{J}\alpha_j = 1,\ \alpha_j \geq \varepsilon,\ \varepsilon \approx 0.1/J$$

Эмпирическая оценка дисперсии:

$$f(\alpha) = n\widehat{\mathrm{Var}}\left(\widehat{\mu}_\alpha\right) + \mu^2 = \sum_{i=1}^{n}\frac{H_{1:J}(x_i)p^2(x_i)}{q_\alpha(x_i)q_{\alpha'}(x_i)} \longrightarrow \min_{\alpha}$$

$$x_i \sim q_{\alpha'} \qquad \left(q_\alpha(x) = \sum_{j=1}^{J}\alpha_j q_j(x)\right)$$

$f(\alpha)$ – выпуклая функция

Множество ограничений можно представить в виде

$x \in S$, $S = \mathrm{conv}(\mathcal{A})$, где A – конечное множество точек

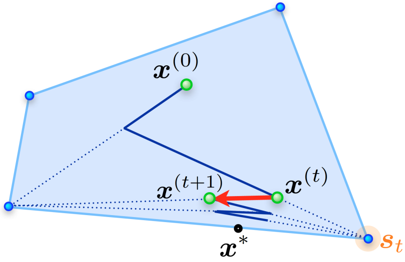
Алгоритм Франка-Вульфа:

$$v_t = \arg\max_{v \in S} \langle v, -\nabla f(\alpha^{(t)}) \rangle$$

$$d_t = v_t - \alpha^{(t)}$$

$$\gamma_t = \frac{2}{t+1}$$

$$\alpha^{(t+1)} = \alpha^{(t)} + \gamma_t d_t$$



Стохастические вариации метода Франка-Вульфа

$$v_t = \arg\max_{v \in S} \langle v, -\widetilde{g}_t \rangle$$

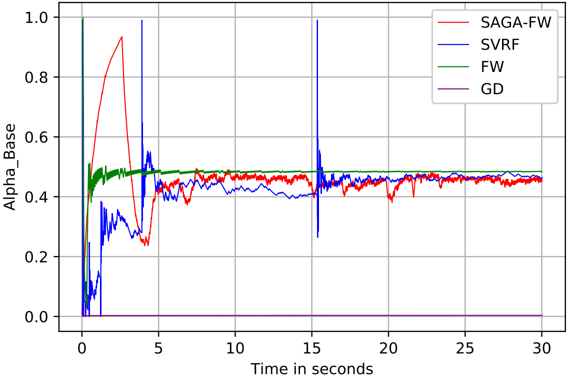
$$\widetilde{g}_t^{SAGA} = \nabla f_j \left(\alpha_j^{(t)} \right) - \nabla f_j \left(\alpha_j^{(t-1)} \right) + \frac{1}{n} \sum_{i=1}^{n} \nabla f_i \left(\alpha_i^{(t-1)} \right)$$

$$\widetilde{g}_t^{SVRF} = \nabla f_j \left(\alpha^{(t)} \right) - \nabla f_j \left(\widetilde{\alpha} \right) + \frac{1}{n} \sum_{i=1}^{n} \nabla f_i \left(\widetilde{\alpha} \right)$$

Эксперименты

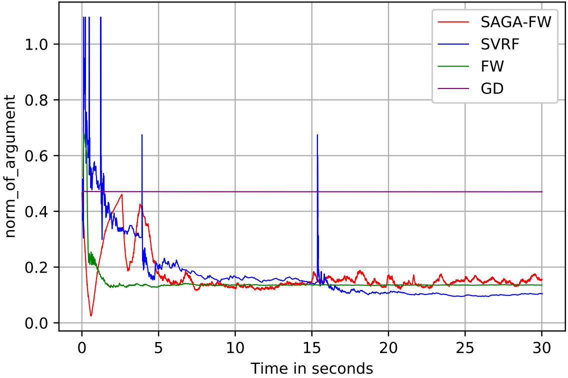
α_{BASE} – вес, с которым для сеплирования выбираются точки из нижнего полупространства

График сходимости по аргументу от времени



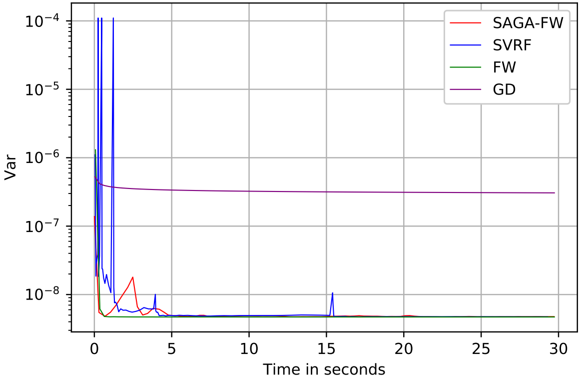
Эксперименты

График сходимости по норме $\|\alpha - \alpha^*\|_2$



Эксперименты

График сходимости по функции от времени



Эксперименты

Метод	SAGA-FW	SVRF	FW	GD
α_{BASE}	0.52	0.48	0.51	0.04
Var	5.08×10^{-9}	5.11×10^{-9}	5.06×10^{-9}	3.59×10^{-7}
cv	1.2	1.2	1.2	9.9

Таблица: Результаты после 10 секунд прогона

$$\mu = \mathrm{P}\left(x \in H\right) = 6.5 \times 10^{-5}$$

Заключение

Спасибо за внимание

Вопросы?