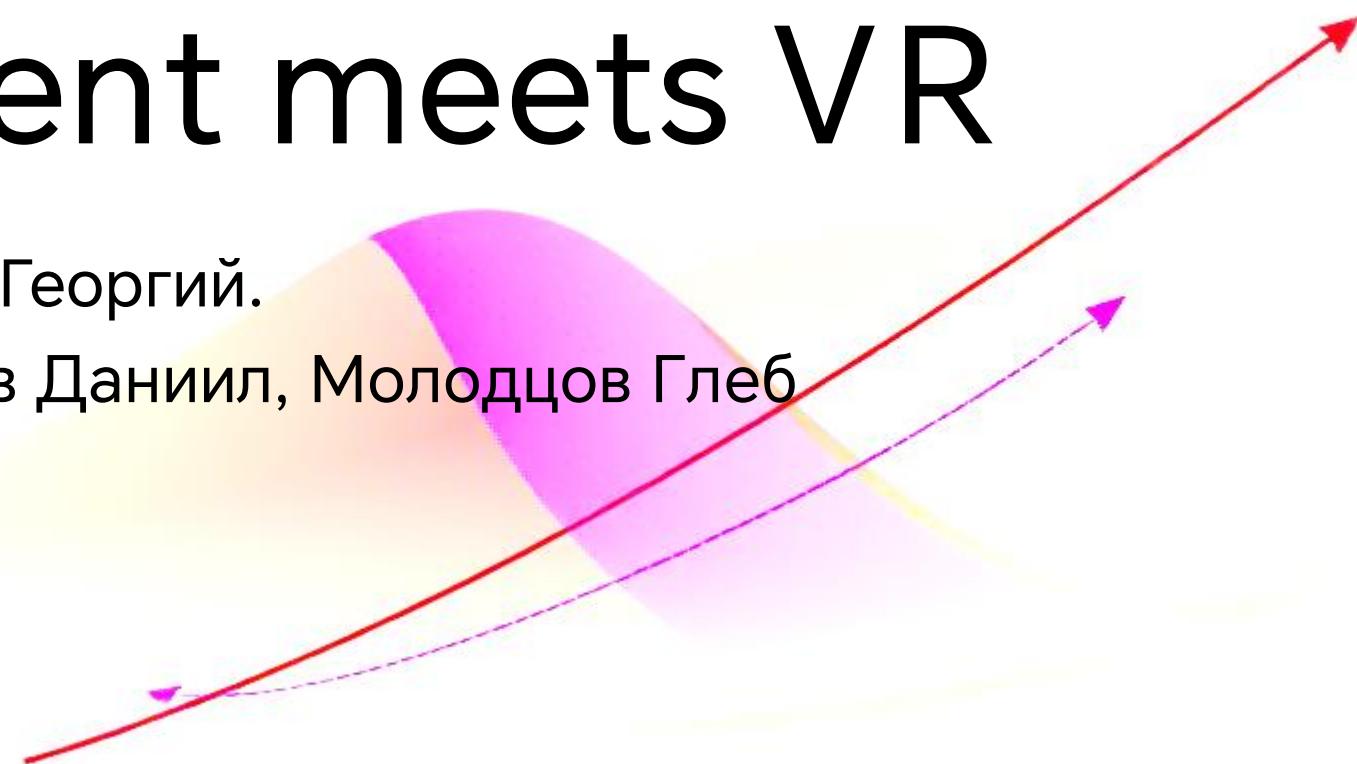


Extragradient meets VR

Кабиков Юрий, Чирков Георгий.

Руководители: Медяков Даниил, Молодцов Глеб

МФТИ, 2025



Мотивация

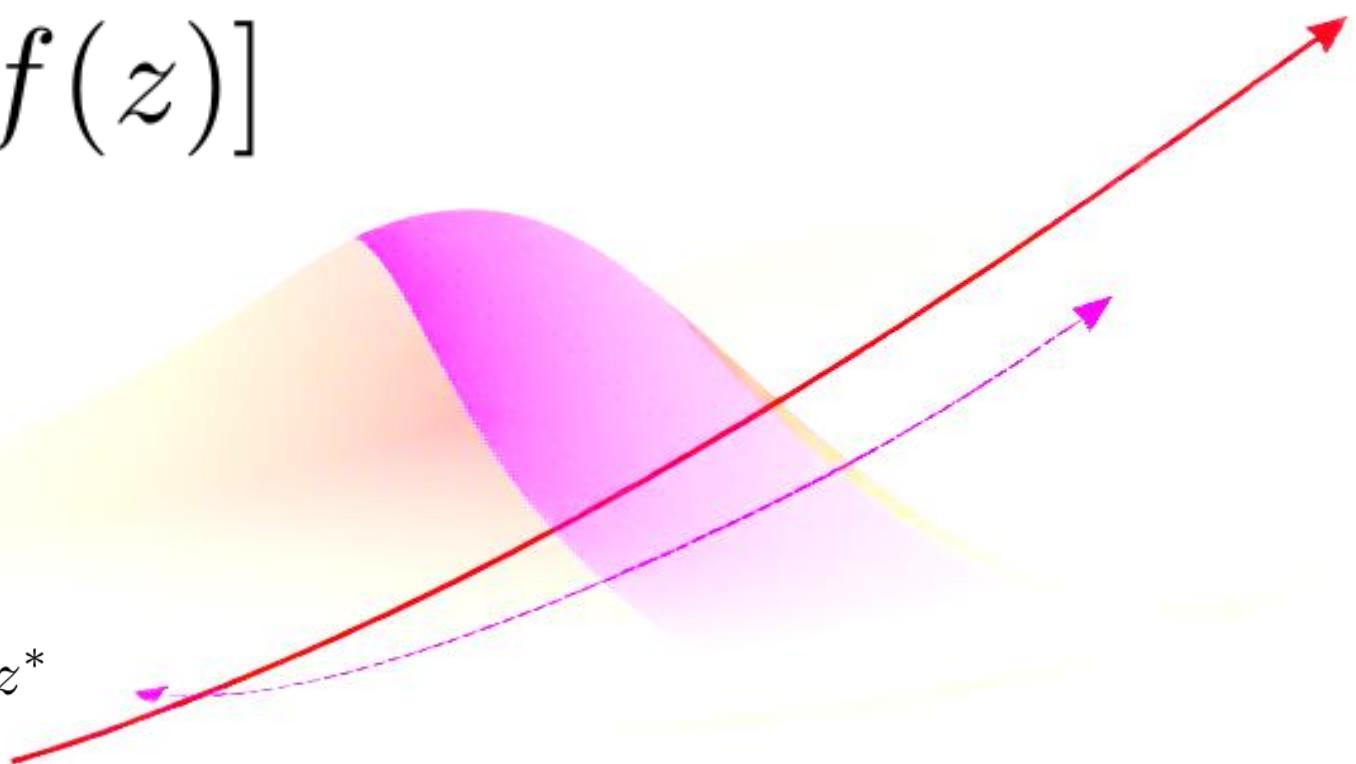
Классическая постановка задачи
(минимизация):

$$\min_{z \in \mathbb{R}^d} [f(z)]$$

Иные задачи:

$$\min_{x \in \mathbb{R}^{d_x}} \max_{y \in \mathbb{R}^{d_y}} [f(x, y)]$$

Find $z^ \in \mathbb{R}^d$ such that $T(z^*) = z^*$*



Вариационные неравенства (VIs)

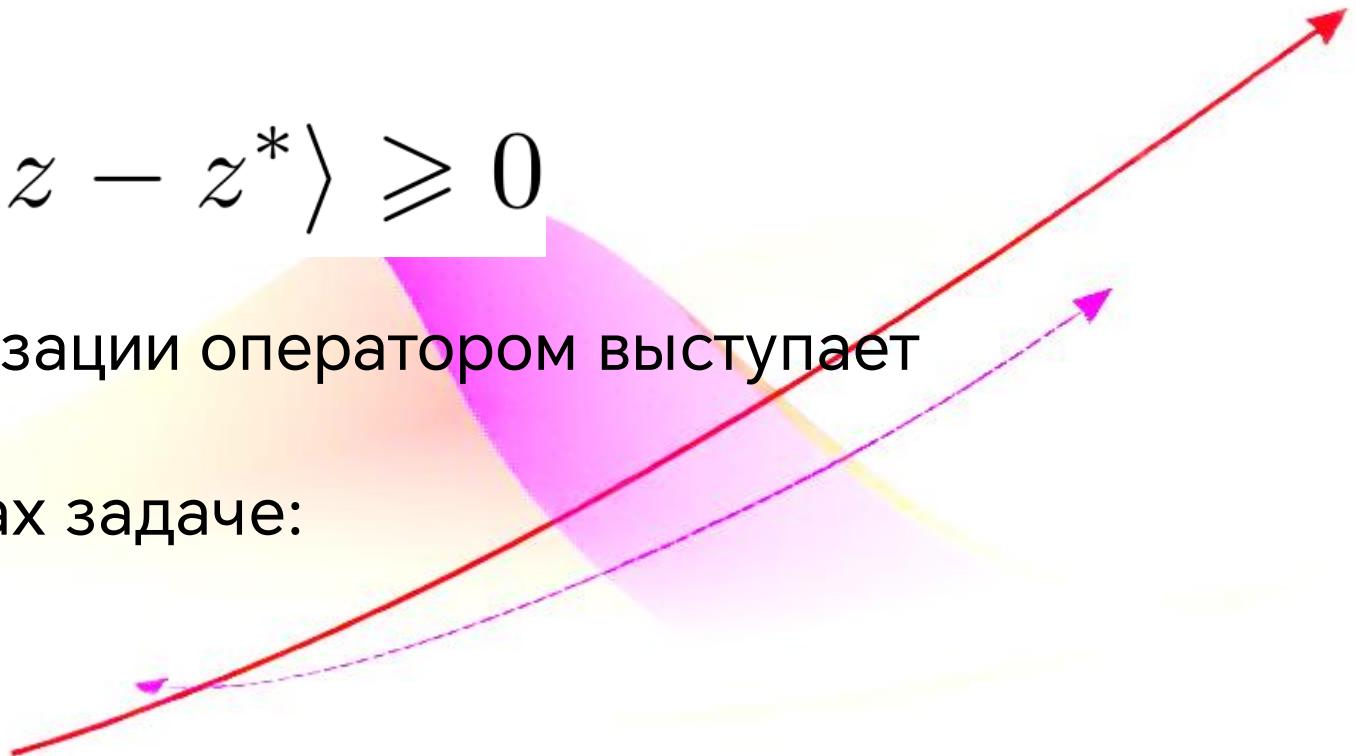
Перечисленные ранее задачи, также, как и другие, могут быть сформулированы в терминах вариационных неравенств:

$$\forall z \in \mathcal{Z} \hookrightarrow \langle F(z^*), z - z^* \rangle \geqslant 0$$

Например, в задаче минимизации оператором выступает градиент целевой функции.

В выпукло-вогнутой min-max задаче:

$$[\nabla_x f(x, y), -\nabla_y f(x, y)]$$



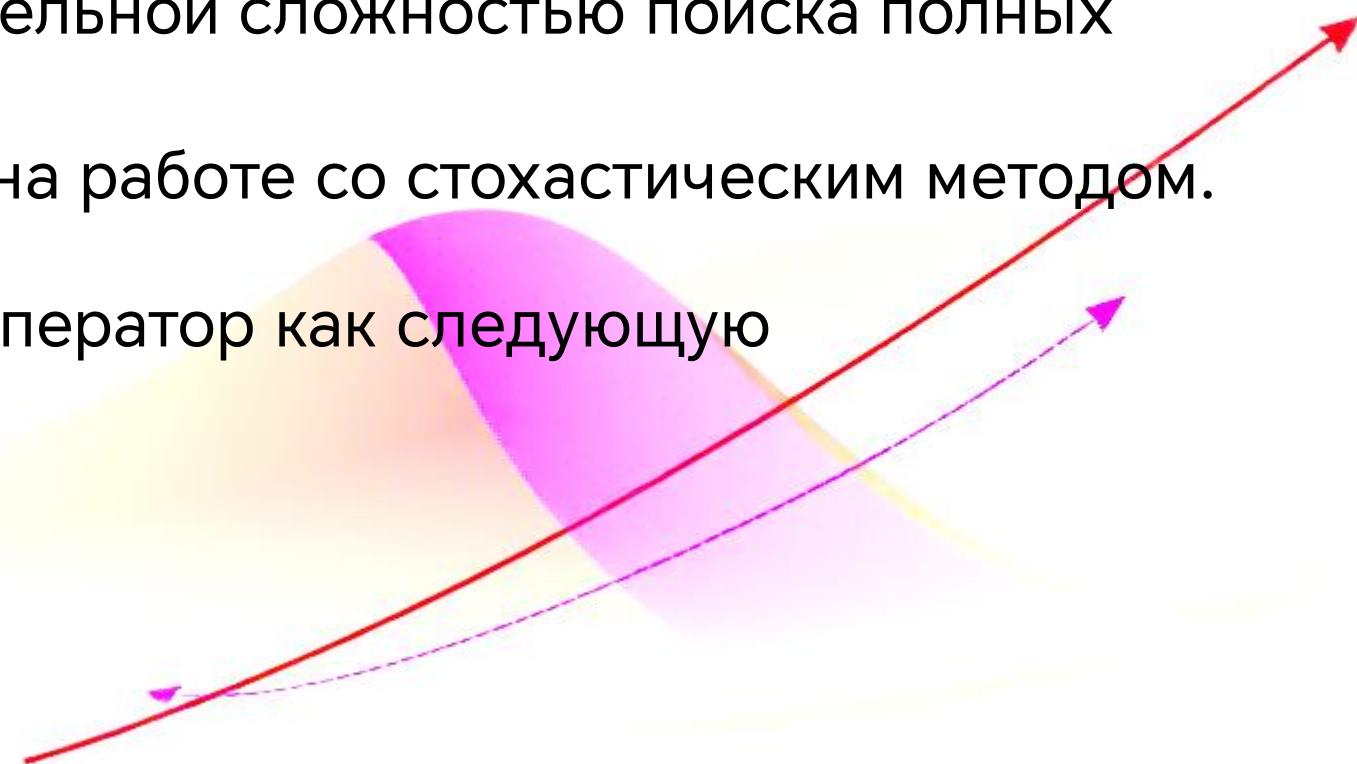
Стохастичность вступает в игру

Широкое применение стохастических методов оптимизации вызвано большой вычислительной сложностью поиска полных операторов.

Наш проект также завязан на работе со стохастическим методом.

Итак, можем рассмотреть оператор как следующую финитную сумму:

$$F(z) = \frac{1}{n} \sum_{i=1}^n F_i(z)$$



Стандартный подход. Stochastic ExtraGradient.

Stochastic ExtraGradient (EG)

- своего рода комбинация

SGD и EG

$$z^{t+\frac{1}{2}} = z^t - \gamma F_{i_t}(z^t),$$

$$z^{t+1} = z^{t+\frac{1}{2}} - \gamma F_{j_t}(z^{t+\frac{1}{2}})$$

Этот метод прост, но содержит существенный недостаток – ненулевую дисперсию:

$$\mathbb{E}\{\|F_{i_t}(z^t) - F(z^t)\|^2\} \not\rightarrow 0, \quad z^t \rightarrow z^*$$

По вышеперечисленной причине внимание приковано к методам, осуществляющим редукцию дисперсии (variance reduction – VR)

ExtraSAGA

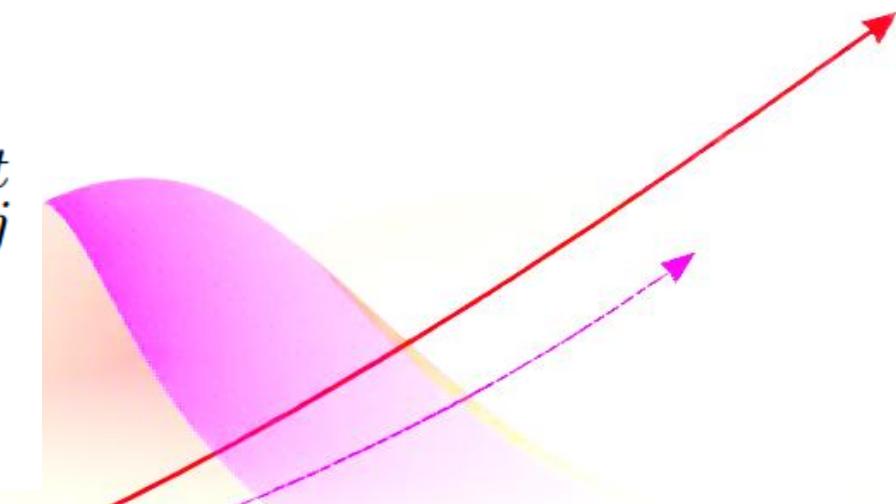
Наш план работы содержит исследование метода, объединившего в себе преимущества методов Optimistic ExtraGradient и SAGA.

$$z^{t+1/2} = z^t - \gamma G^{t-1}$$

$$G^t = F_{i_t}(z^{t+1/2}) - y_{i_t}^t + \frac{1}{n} \sum_{j=1}^n y_j^t$$

$$y_i^{t+1} = F_{i_t}(z^{t+1/2}) \text{ for } i = i_t$$

$$z^{t+1/2} = z^t - \gamma G^t$$



Благодаря использованию несмещенного аппроксиматора оператора из метода SAGA, его дисперсия стремится к нулю.

На данный момент проведено

Изучение избранных стохастических методов оптимизации, различных постановок задач. В частности, проработка методов ExtraGradient и SAGA в отдельности.

Тщательное ознакомление с результатами похожей по смыслу работы руководителей проекта о совмещении EG и PAGE в методе ExtraPAGE.

Доказательство редукции дисперсии в исследуемом методе

Предложение вариантов модернизации метода путём использования другого способа перехода в промежуточную точку

Дальнейшие планы

Окончательное доказательство сходимости метода (или его одной из модернизированных версий) при хорошей скорости сходимости

Эксперименты, тестирование метода (GANs, Adversarial training, Denoising, min-max)

Изучение ML для более тщательного понимания применимости исследуемого метода

Написание научной статьи