

. Вопросы по докладу «Натуральный градиент и стохастическая оптимизация»

- 1) Запишите формулу для расстояния между векторами θ и $\theta + \delta\theta$ в римановом пространстве. Запишите оптимизационную задачу для вывода формулы натурального градиента и ее решение.
- 2) Приведите пример оптимизационной задачи, когда переход к новой параметризации и использование натурального градиента увеличивают скорость сходимости градиентного спуска
- 3) Какая существует взаимосвязь между дивергенцией Кульбака-Лейблера и информационной матрицей Фишера?
- 4) В чем проблема использования обычного метода стохастического градиентного спуска? Опишите принцип работы метода Momentum.
- 5) Запишите формулы методов спуска Adagrad и RMSProp. В чем заключается главное различие между методами?
- 6) Запишите примерный алгоритм работы метода стохастического спуска Adam.