

62-я научная конференция МФТИ  
18–24 ноября 2019 года, Москва



## Метод онлайн-обучения для нейросетевого управления космическим аппаратом с двигателем малой тяги

М.Г. Ширококов<sup>1</sup> С.П. Трофимов<sup>1</sup> А.В. Сорокин<sup>2</sup>

<sup>1</sup>Институт прикладной математики им. М.В. Келдыша РАН

<sup>2</sup>Московский физико-технический институт

Работа поддержана грантом РФФИ №18-31-00403

# Содержание

- ▶ Цель работы
- ▶ Постановка задачи
- ▶ Рекуррентные нейронные сети
- ▶ Решение задачи
- ▶ Моделирование
- ▶ Заключение

# Цель работы

- ▶ Применить и протестировать метод онлайн-обучения для управления космическим аппаратом (КА), снабженным двигателем малой тяги, в задаче оптимального многовиткового перелета между точками пространства.

# Постановка задачи

- ▶ Рассматриваются движения КА в околоземном пространстве в рамках двух моделей: простой (ньютоново поле притяжение Земли) и точной (поле притяжение Земли, Луны и Солнца).
- ▶ Рассматривается задача оптимального перелета из точки  $\mathbf{x}_1$  в точку  $\mathbf{x}_2$  за фиксированное время  $T$ ;  $\mathbf{x}_1 = [\mathbf{r}_1, \mathbf{v}_1]$ ,  $\mathbf{x}_2 = [\mathbf{r}_2, \mathbf{v}_2]$ . Под оптимальностью понимается минимизация функционала  $\int_0^T u^2 dt \rightarrow \min$ . Перелет между геопереходной орбитой и геостационарной орбитой.
- ▶ В рамках простой модели создается и обучается нейронная сеть, прогнозирующая движение аппарата с малой тягой. В рамках сложной модели (в режиме полета) эта сеть адаптируется с помощью метода онлайн-обучения.
- ▶ В качестве архитектуры сети выбрать рекуррентные сети.

# Рекуррентные нейронные сети

Рекуррентные нейронные сети (Recurrent Neural Networks, RNN) — это нейронные сети с одной или несколькими обратными связями (слоями, замкнутыми на себя). Обратные связи дают возможность обрабатывать серии событий во времени или последовательные пространственные цепочки. Обладают внутренней памятью.

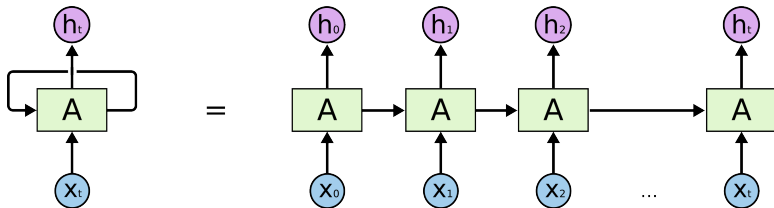


Рис. 1: RNN и ее развернутое представление.

# План решения задачи

- ▶ Ввести сопряженные переменные (принцип максимума Понтрягина)
- ▶ Инициализировать обучающие выборки
- ▶ Обучение нейронных сетей для прогнозирования и управления движением
- ▶ Обучение сетей в рамках простой модели
- ▶ Применение метода онлайн-обучения прогнозирующей нейронной сети для адаптации к сложной модели движения аппарата
- ▶ Рассмотреть возможность онлайн-обучения управляющей нейронной сети аппарата к сложной модели

# Выборка для прогнозирующей сети

Сначала в некоторой  $\varepsilon$ -окрестности начальной точки  $\mathbf{s}_1 = [\mathbf{r}_1, \mathbf{v}_1, \boldsymbol{\lambda}_r(0), \boldsymbol{\lambda}(0)]$  расширенного пространства случайным образом выбираются точки, определяющие трубку траекторий с точками на ней:  $S_i = \{[\mathbf{r}_n^i, \mathbf{v}_n^i, \boldsymbol{\lambda}_{r_n^i}, \boldsymbol{\lambda}_{v_n^i}, t_n^i, \mathbf{r}_{end}^i]\}_{n=1}^N$

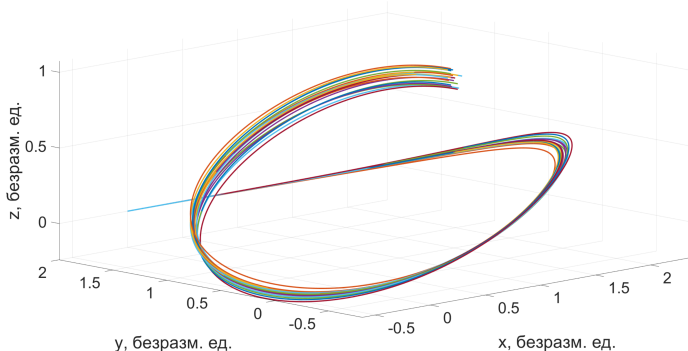


Рис. 2: Пример трубки траекторий.

# Выборка для управляющей сети

Вокруг номинальной траектории в фазовом (нерасширенном) пространстве случайным образом выбираются точки, им сопоставляются значения сопряженных переменных, чтобы попасть в цель  $\mathbf{x}_2$ :  $S_{control} = \{[r_i, \mathbf{v}_i, t_i, \lambda_{r_i}, \lambda_{\mathbf{v}_i}]\}_{i=1}^N$

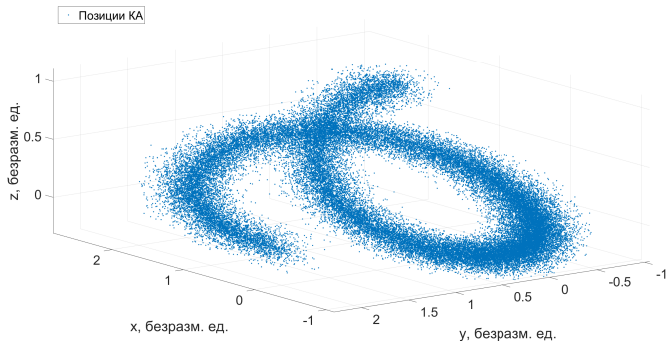


Рис. 3: Визуализация точек, генерирующих обучающую выборку.



# Метод онлайн-обучения

- ▶ Прогнозирующая сеть была предварительно обучена с помощью стохастического градиентного спуска. Два слоя: первый – рекуррентный на 512 нейронов (длина последовательности – 30) и второй слой на 64 нейронов.
- ▶ В обеих моделях оптимизируемый функционал – квадрат невязки:  $F = (\mathbf{x}_{pred} - \mathbf{x}_2)^T (\mathbf{x}_{pred} - \mathbf{x}_2)$ , где  $\mathbf{x}_{pred}$  - прогнозируемый сетью вектор,  $\mathbf{x}_2$  - целевой вектор

$$\frac{\partial F(\mathbf{x}_{pred}(\omega))}{\partial \omega_i} = 2(\mathbf{x}_{pred} - \mathbf{x}_2)^T \frac{\partial \mathbf{x}_{pred}}{\partial \omega_i}$$

$$\Delta \omega = -\mathbf{J}^T \underbrace{(\mathbf{J}\mathbf{J}^T + \lambda)}_{\text{скаляр}}^{-1} F, \quad \mathbf{J} = \frac{\partial F}{\partial \omega}$$

- ▶ Метод не требует обращения матриц, что удобно для использования на борту КА

# Моделирование онлайн-обучения

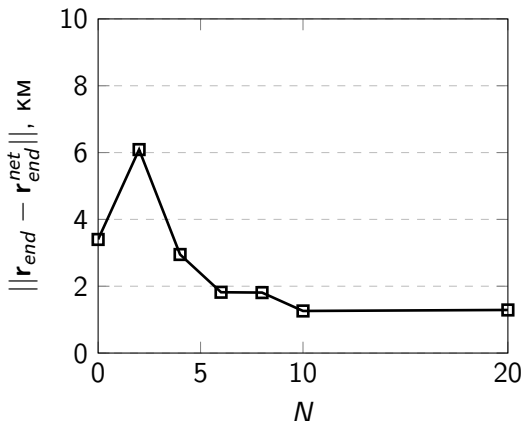


Рис. 4: Точность попадания в целевую точку в зависимости от числа итераций в режиме онлайн-обучения.

# Онлайн-обучения для управляющей сети

- ▶ Проблема: сопряженные переменные недоступны измерениям и не были определены в рамках сложной модели
- ▶ Возможное решение: объединить нейронные сети – на вход прогнозирующей сети подать выход управляющей сети
- ▶ Изменение весов единой нейронной сети

$$Net_{pred}(\mathbf{x}, Net_{control}(\mathbf{x})) = \mathbf{x}_{pred}$$

# Заключение

- ▶ В задаче перелета между геопереходной и геостационарной орбитами с малой тягой введены прогнозирующая и управляющая нейронные сети с рекуррентной архитектурой.
- ▶ Метод онлайн-обучения успешно применен для адаптации прогнозирующей сети к сложной модели. На примере показано, что всего за 10 итераций невязка за счет адаптации уменьшается почти в два раза.
- ▶ Предложен метод онлайн-обучения и для управляющей нейронной сети, основанный на идее объединения нейронных сетей.