

# Применение нейронных сетей для моделирования объектов управления

Сивков Антон Александрович

научный руководитель: д.ф.-м.н., профессор Фомичев В. В.

Кафедра нелинейных динамических систем и процессов управления  
Факультет вычислительной математики и кибернетики  
Московский государственный университет имени М.В. Ломоносова

2019

# Постановка задачи

- ▶ Решается задача идентификации
- ▶ Объект моделируется как 'чёрный ящик'
- ▶ Сеть обучается предсказывать выходные сигналы объекта управления по входным
- ▶ Сравнивается способность сетей с различными архитектурами экстраполировать поведение объекта

Объект управления, использованный для моделирования:

$$H = \frac{z}{(z - \frac{1}{2})(z - \frac{1}{3})(z - \frac{1}{4})}, \quad dt = 1 \text{ sec.}, \quad x_0 = 0 \quad (1)$$

# Многослойный перцептрон

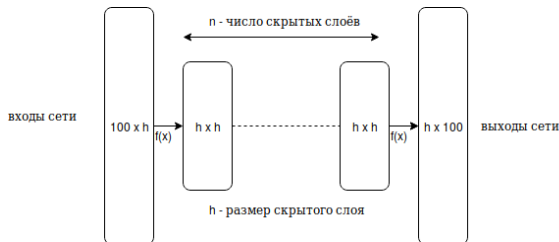
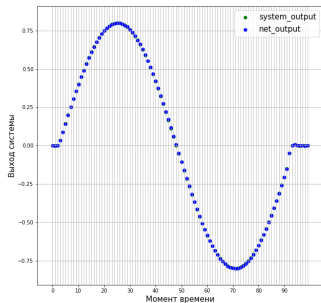


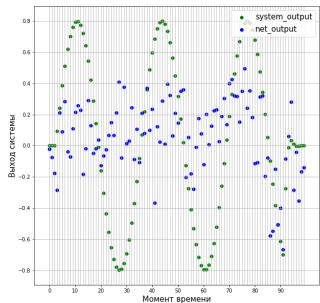
Рис.: Схема многослойного перцептрона

Размер скрытого слоя  $h$ , число скрытых слоев  $n$ , функция активации  $f(x)$  — гиперпараметры

# Эксперимент с многослойным перцептроном



(a) Предсказание при отсутствии сдвига частот



(b) Предсказание при сдвиге частоты на 0.1 Гц (частота 0.2 Гц)

Рис.: Предсказание поведения объекта

# Выводы для многослойного перцептрона

- ▶ Сеть неспособна экстраполировать поведение объекта. Среднеквадратичная ошибка при сдвиге частот  $\approx 0.2$ . Перебор гиперпараметров сети не даёт существенных улучшений.
- ▶ В статьях по теме идентификации систем с помощью многослойного перцептрона решают задачи идентификации для сложных нелинейных объектов без тестирования способности сети к экстраполяции.

# Сеть с долгосрочно–краткосрочной памятью (LSTM)

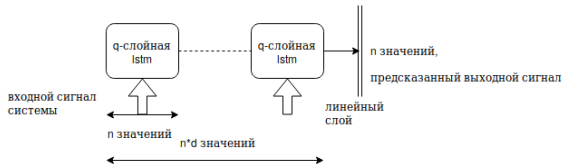
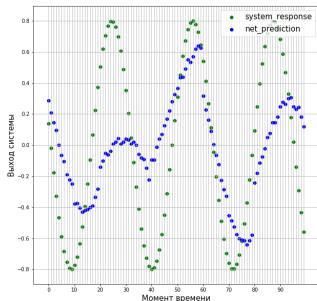


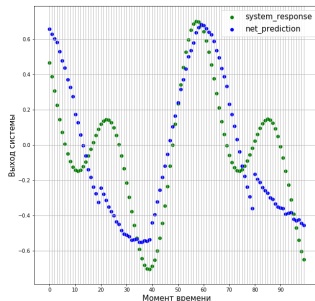
Рис.: Схема основанной на lstm ячейке сети

Размер шага рекурсии, входа и выхода ячейки  $n$ , максимальная глубина рекурсии  $d$ , число внутренних слоев ячейки  $q$  — гиперпараметры.

# Эксперимент с lstm сетью



(a) Предсказание при сдвиге частоты на 0.1 Гц (частота 0.2 Гц)



(b) Предсказание при подаче на вход сети смеси сигналов с частотой 0.1 Гц и 0.2 Гц

Рис.: Предсказание поведения объекта

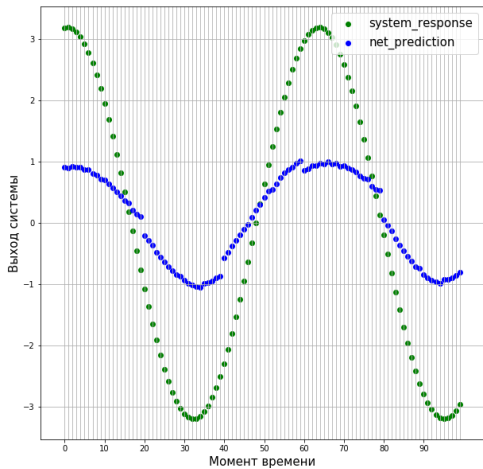


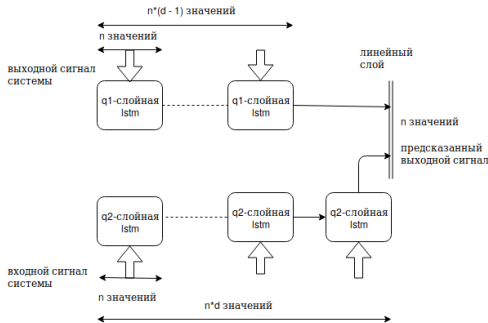
Рис.: Предсказание при увеличенной амплитуде сигнала



## Выводы для lstm сети

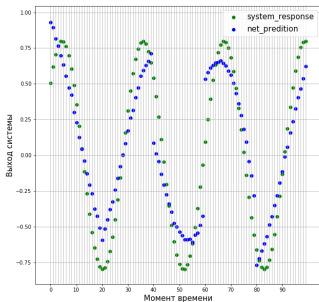
- ▶ Сеть допускает значительные ошибки, однако среднеквадратичная ошибка предсказания уменьшается на порядок по сравнению с многослойным перцептроном. Среднеквадратичная ошибка при сдвиге частот  $\approx 0.02$
- ▶ Интересен тот факт, что с наращиванием сложности сети с помощью гиперпараметров, ошибка на смещённой выборке **возрастает практически монотонно**.
- ▶ Идея использования lstm для идентификации встречается в статьях, однако исследования способности сетей к экстраполяции затруднительно найти.

# lstm сеть, использующая данные выхода системы для предсказания

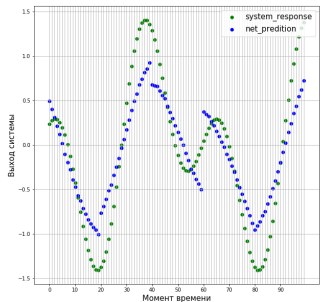


**Рис.:** Схема основанной на lstm ячейке сети, использующей выходной сигнал системы

Размер шага рекурсии, входа и выхода ячейки  $n$ , максимальная глубина рекурсии  $d$ , количество внутренних слоев ячеек  $q1$  и  $q2$  — гиперпараметры.



(a) Предсказание при сдвиге частоты на 0.1 Гц (частота 0.2 Гц)



(b) Предсказание при подаче на вход сети смеси сигналов с частотой 0.1 Гц и 0.2 Гц

Рис.: Предсказание поведения объекта

## Выводы для улучшенной lstm сети

- ▶ Сеть, использующая собственные предсказания на предыдущих моментах времени как входные данные, предсказывает выход существенно лучше, чем сеть, использующая только входные сигналы.  
Среднеквадратичная ошибка при сдвиге частот  $\approx 0.012$ .
- ▶ Наблюдается аналогичный предыдущему случаю эффект возрастания ошибки на смещенной выборке при росте сложности сети.
- ▶ Идея использования предсказаний в предыдущие моменты времени для следующих встречается в одной из тематических статей, однако такой подход широко используется в различных задачах, решаемых с помощью нейронных сетей.

# Выводы

- ▶ Проведен сравнительный анализ различных архитектур сетей, однозначно выявлены сети, обладающие лучшей способности к экстраполяции поведения объекта управления.
- ▶ Определен примерный вектор дальнейшего поиска наиболее адекватной задаче архитектуры сети, а именно: сеть должна быть рекуррентной; должна использовать предсказания выхода объекта управления в предыдущие моменты времени как входы; не должна обладать чрезмерным количеством параметров во избежание переобучения.

Спасибо за внимание