Байесовская дистилляция моделей на базе трансформеров

A Preprint

Игорь Н. Игнашин*
Кафедра интеллектуальныз систем
Национальный исследовательский университет «Московский физико-технический институт»
Россия, 141701, г. Долгопрудный, Институтский пер., д. 9
ignashin.in@phystech.edu

Elias D. Striatum
Department of Electrical Engineering
Mount-Sheikh University
Santa Narimana, Levand
stariate@ee.mount-sheikh.edu

15.12.2023

В данной работе исследовано несколько способов дистилляции моделей на базе трансформеров, а также модели RNN. В статье в первую очередь уделяется внимание дистилляции посредством удаления слоя Attention из моделей. В качестве базовых моделей для дистилляции взяты модель RNN с аддитивным вниманием и модель трансформера seq2seq для задачи перевода текстов. Получены априорные распределения для инициализации дистиллированной модели RNN.

Ожидается лучшая сходимость, лучшее качество у дистиллированной модели по сравнению с моделью той же структуры, но с параметрами инициализации из произвольного нормального распределения при одинаковых условиях обучения.

Keywords Байесовская дистилляция · Трансформеры · RNN

1 Введение

Модели трансформеров и RNN часто используются во многих сложных задачах машинного обучения, в том числе, машинного перевода. Однако иногда модели являются переусложненными для своих задач. Потому возникает проблема уменьшения размерности этих моделей или дистилляции моделей. В данной статье поднимается вопрос дистилляции моделей посредством удаления слоев внимания. В статье ? разбирается дистилляция моделей за счет удаления линейных слоев. Подобно тому, как происходит дистилляция модели за счет удаления слоев в многослойной неросети, можно удалять слои внимания в трансформерах, использующих большое количество слоев внимания.

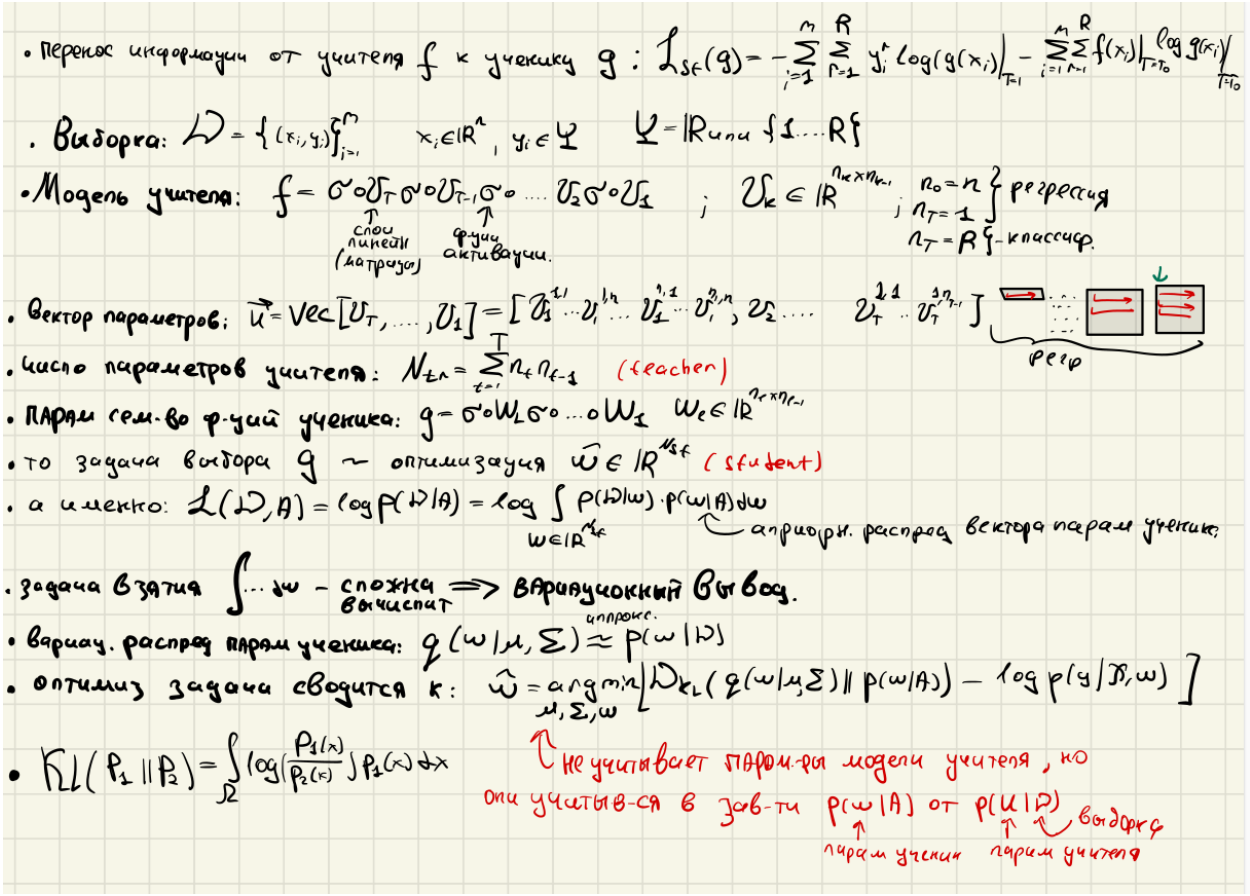

Рис. 1: Sample figure caption.

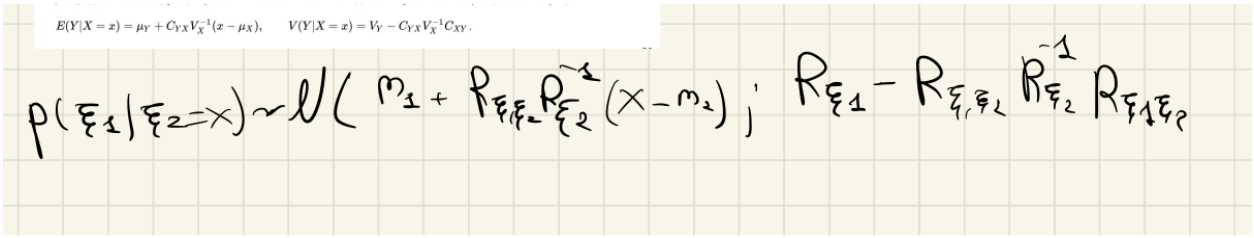

Рис. 2: Sample figure caption.

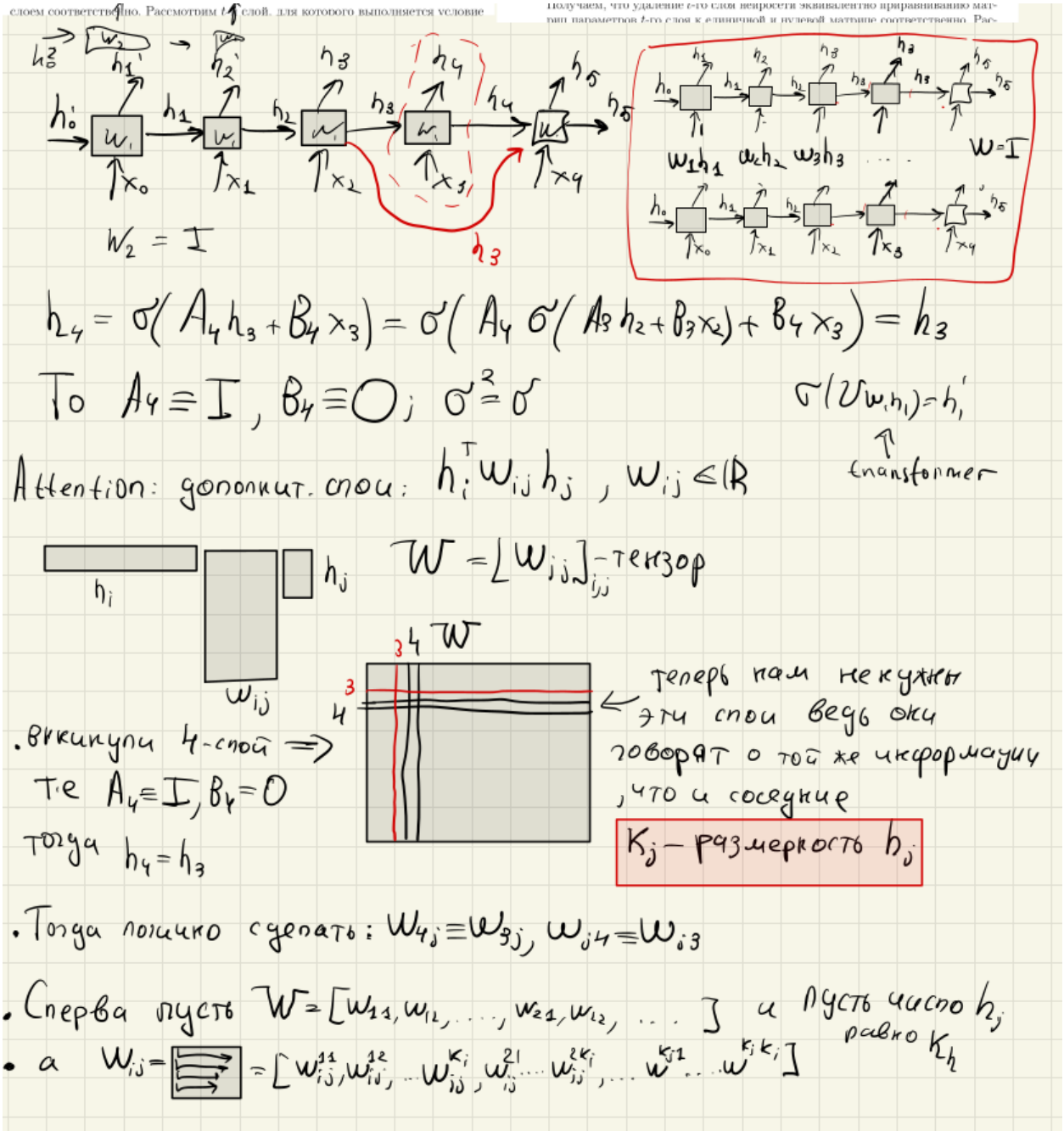

Рис. 3: Sample figure caption.

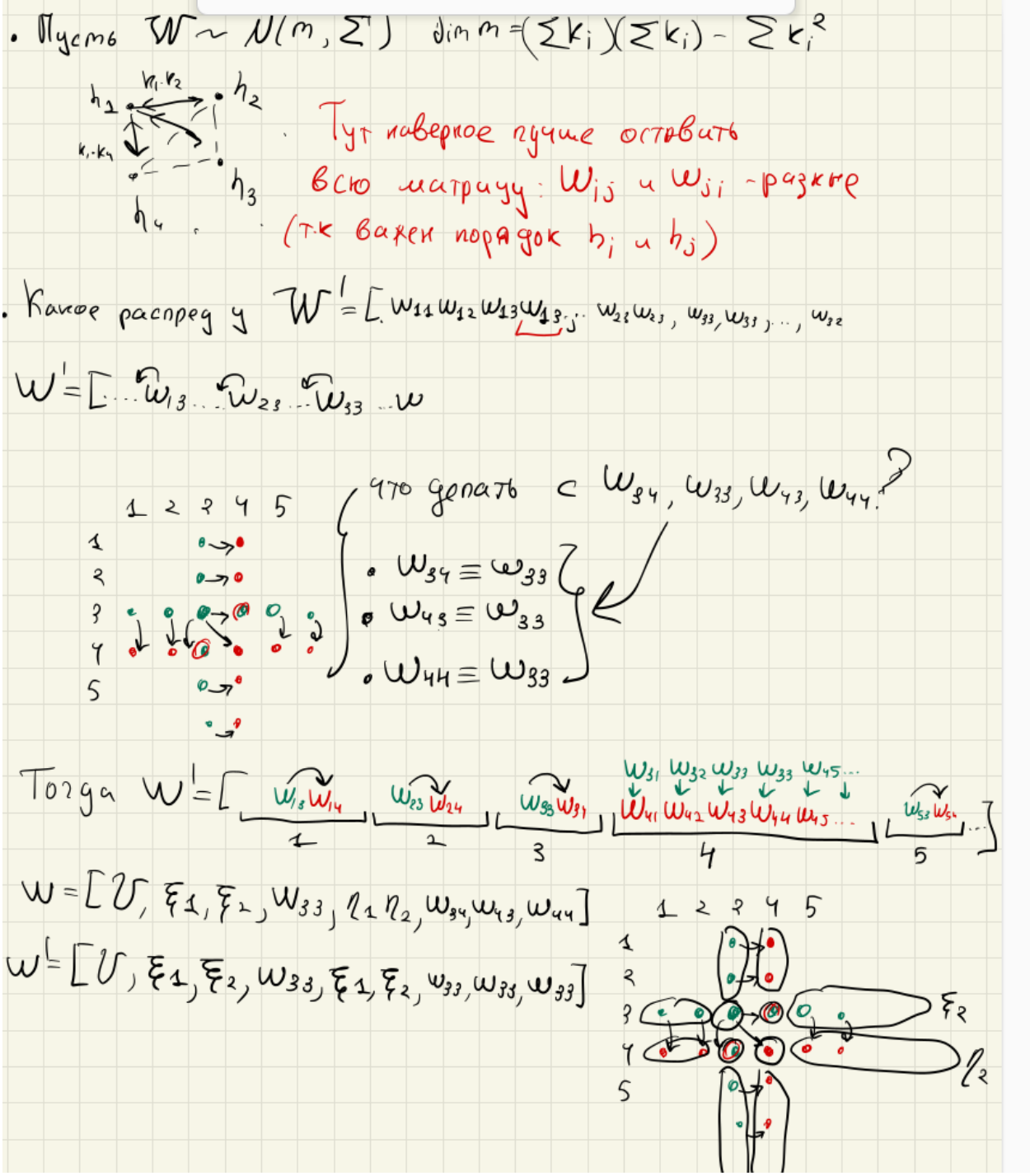

Рис. 4: Sample figure caption.

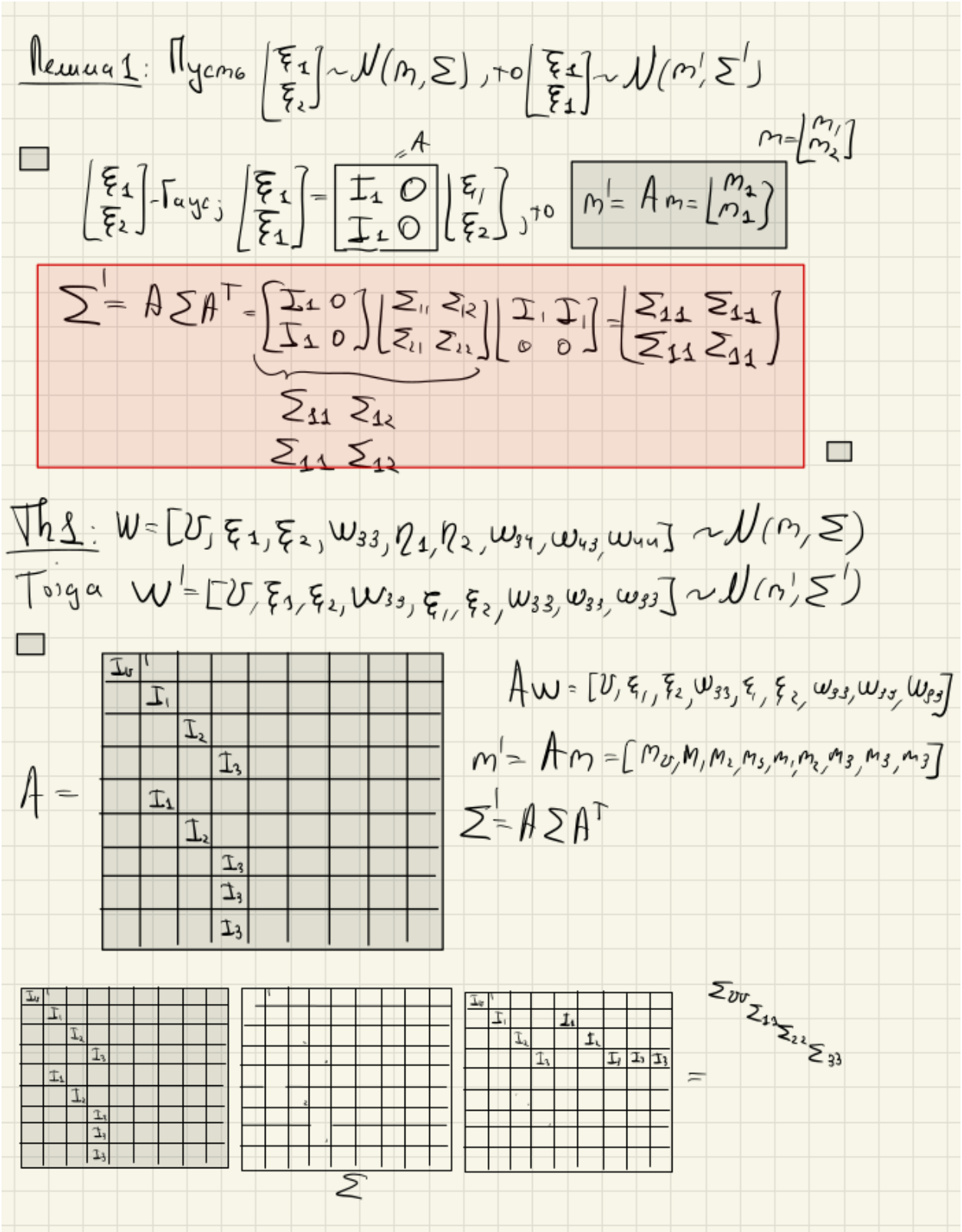

Рис. 5: Sample figure caption.

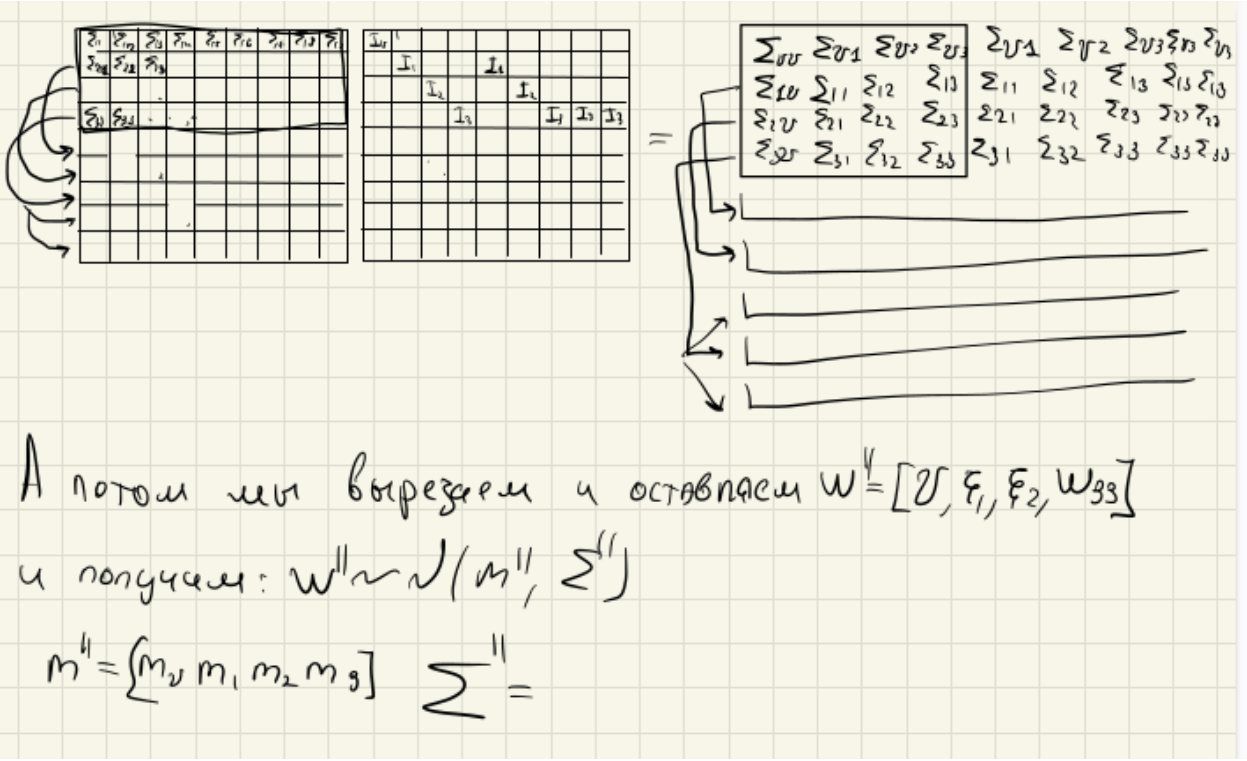

Рис. 6: Sample figure caption.

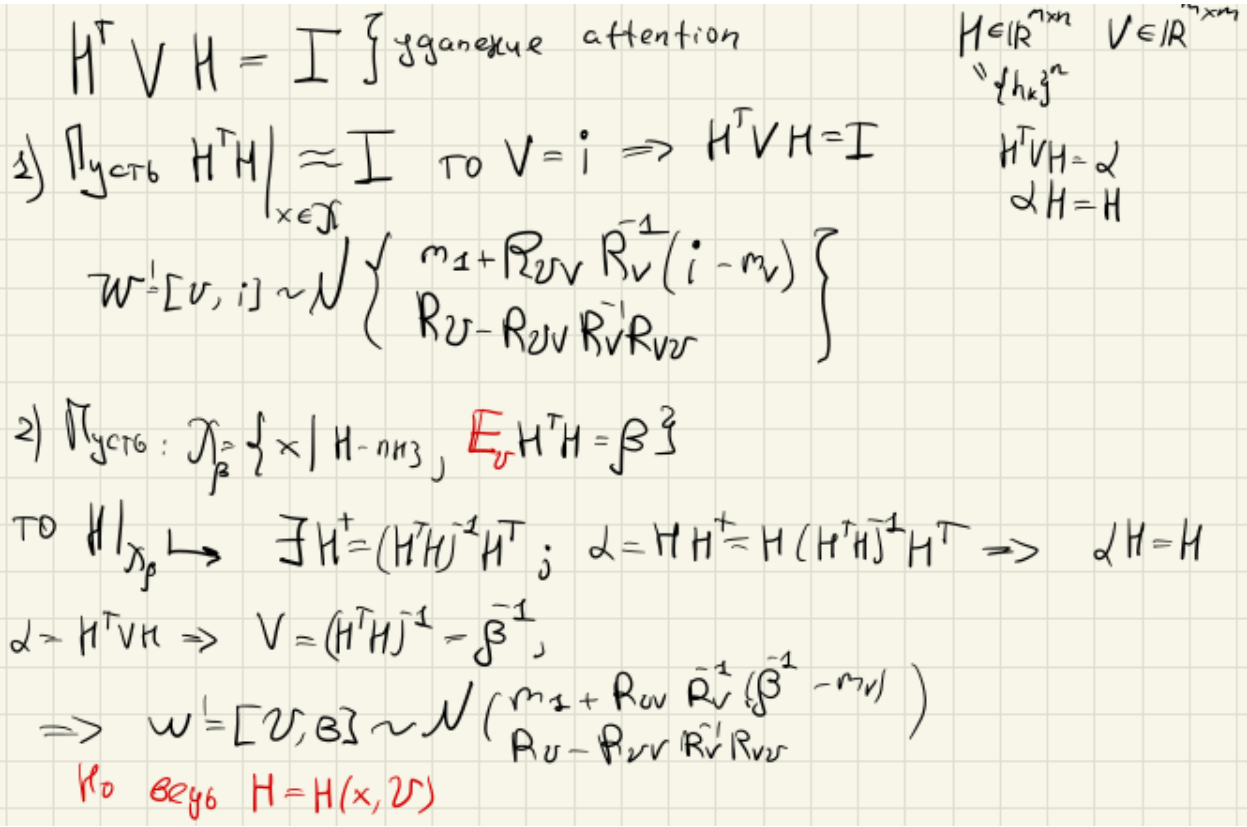

Рис. 7: Sample figure caption.

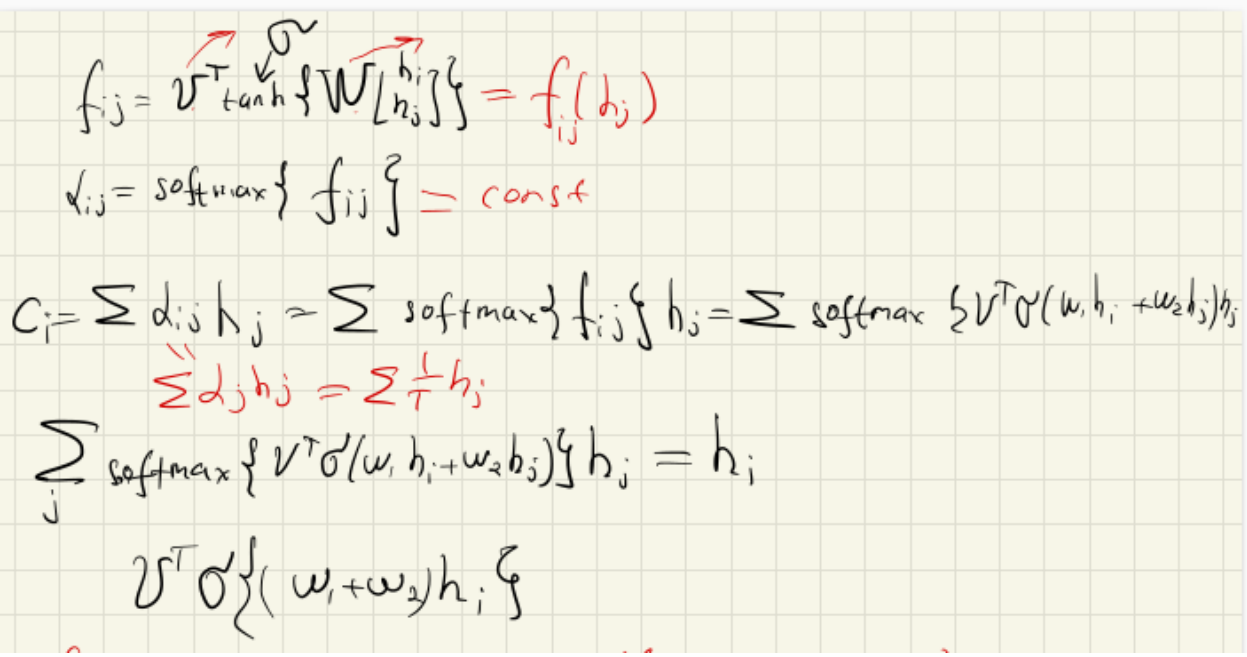

Рис. 8: Sample figure caption.

2 Постановка задачи

3 Дистилляция

3.1 Трансформер

3.2 RNN

Список литературы

*

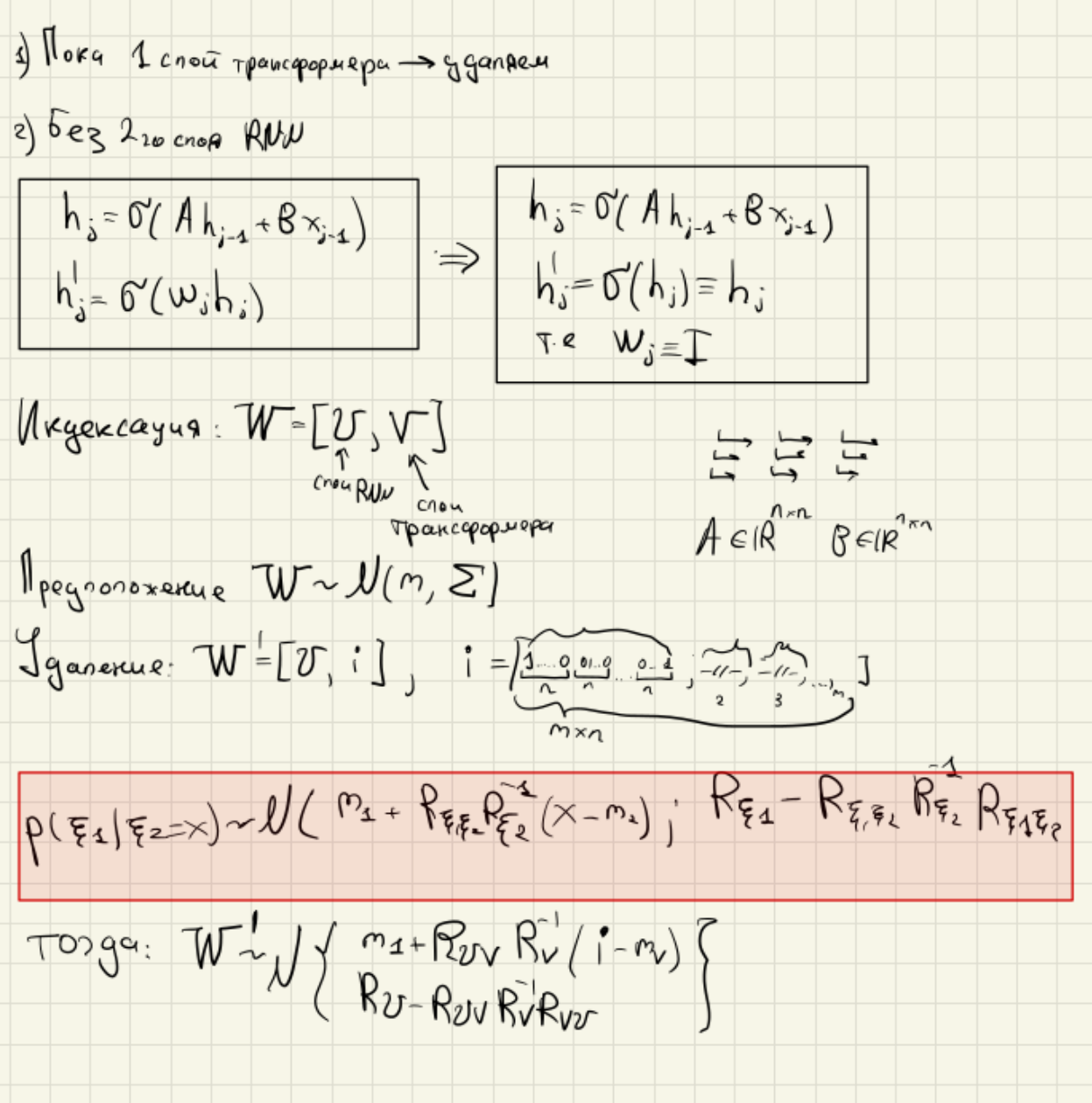

Рис. 9: Sample figure caption.

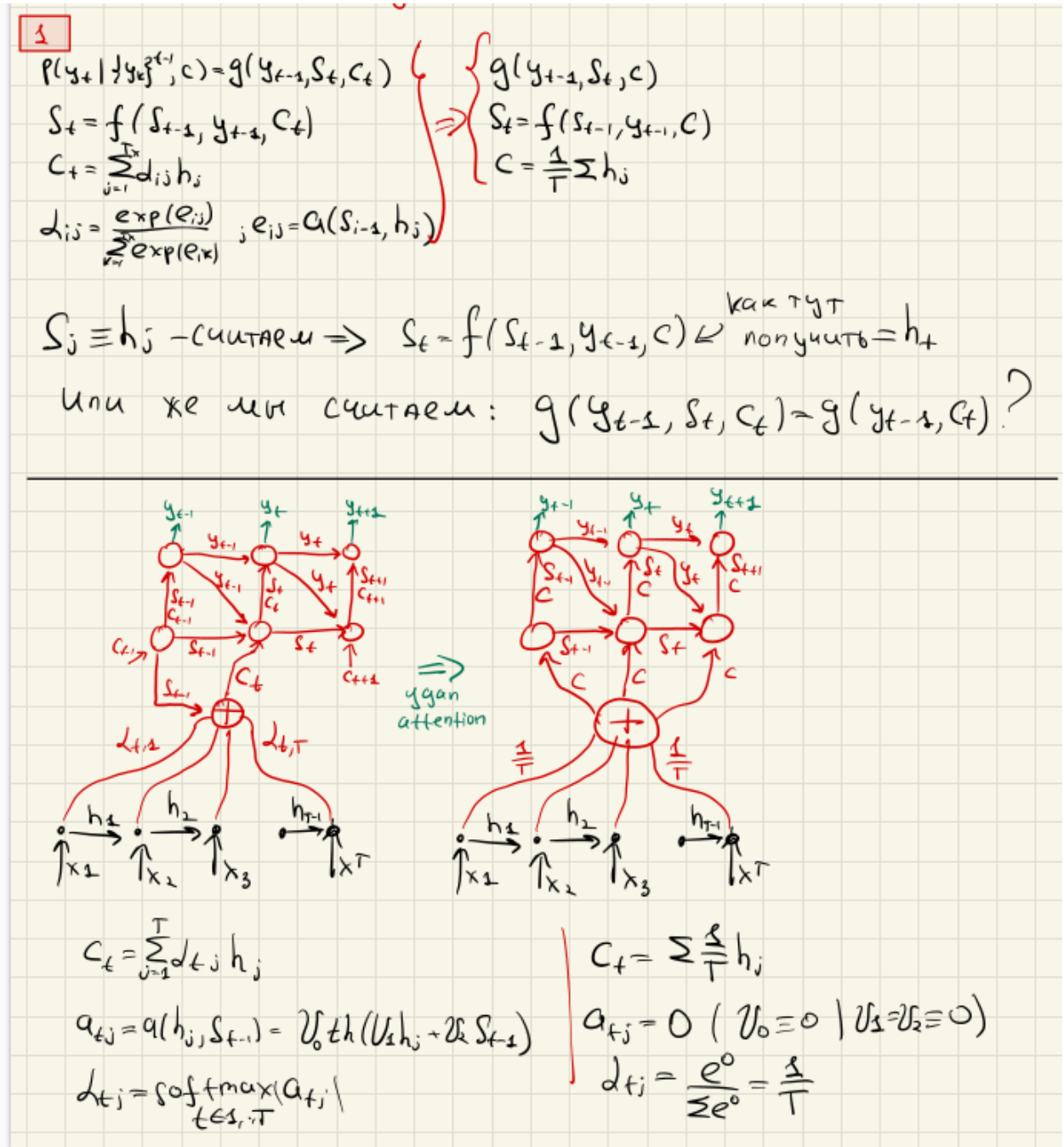

Рис. 10: Sample figure caption.

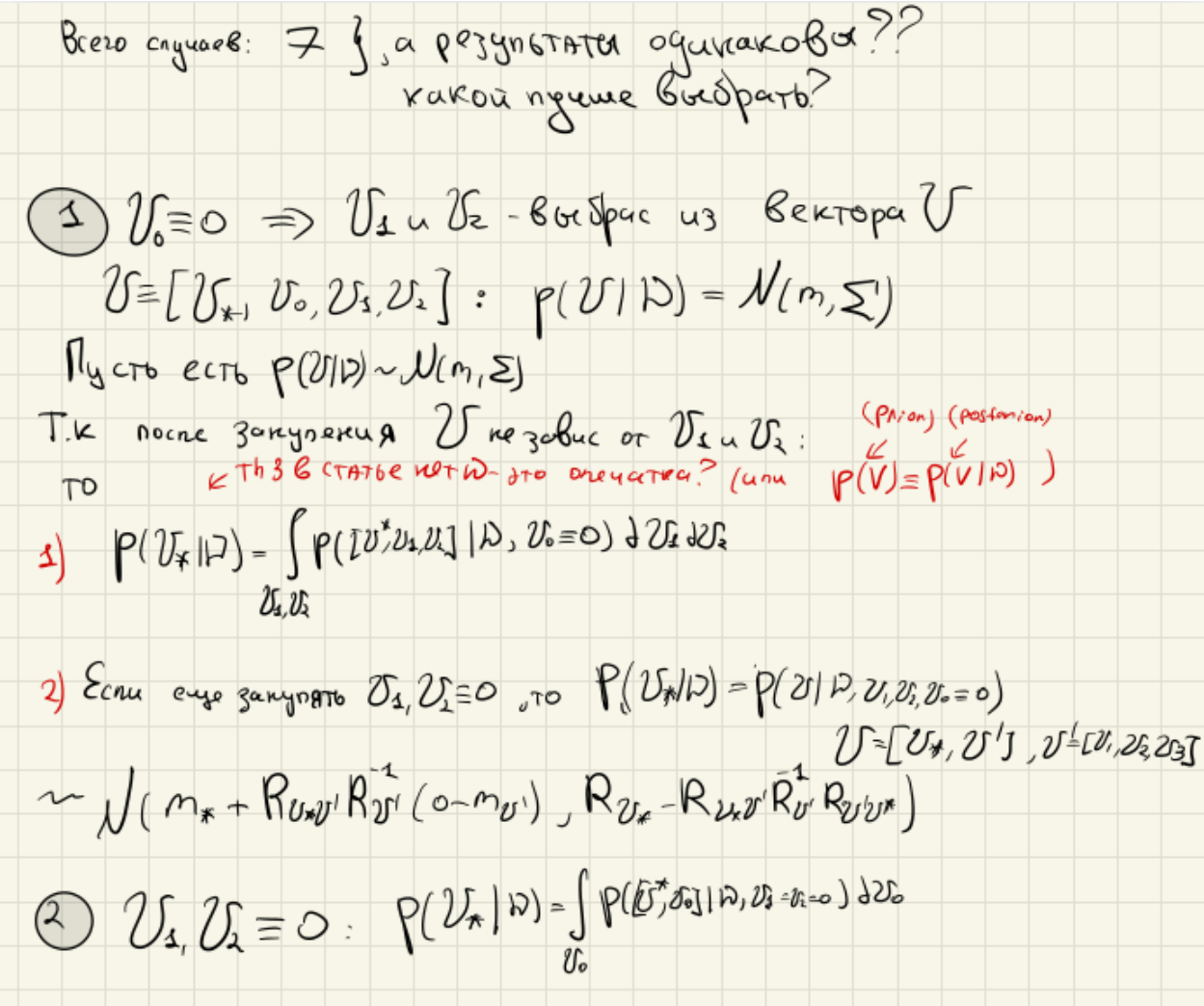

Рис. 11: Sample figure caption.