

# Обучение и вывод в модели Ограниченной Машины Больцмана

Хомутов Н.Ю.

Научная группа Байесовских методов  
ВМК МГУ им. М.В. Ломоносова

13 мая 2014 г.

## 1 Модель ограниченной машины Больцмана

## 2 Задача вывода

## 3 Задача обучения

## 4 Эксперименты

- Архитектура
- Оценка градиента
- Регуляризаторы
- Вывод

# Ограниченная машина Больцмана

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

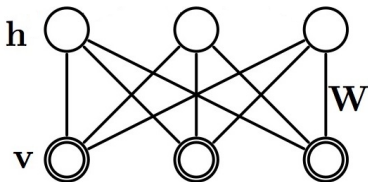
Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Restricted Boltzmann Machine – марковское случайное поле:

- Переменные – бинарные. Принимают значения 0 и 1
- Несмежные переменные условно независимы
- Есть слой видимых переменных, один слой скрытых
- Связи есть только между слоями. Внутренние связи отсутствуют



# Ограниченная машина Больцмана

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Restricted Boltzmann Machine задаёт совместное распределение на переменные  $\mathbf{v}$  и  $\mathbf{h}$

$$P(\mathbf{v}, \mathbf{h} | \Theta) = \frac{1}{Z(\Theta)} \exp(-E(\mathbf{v}, \mathbf{h}, \Theta))$$

$$E(\mathbf{v}, \mathbf{h}, \Theta) = -\mathbf{b}_v^T \mathbf{v} - \mathbf{b}_h^T \mathbf{h} - \mathbf{v}^T W \mathbf{h}$$

Здесь:

- $\mathbf{v} = (v_1, \dots, v_n)^T$  – вектор видимых переменных
- $\mathbf{h} = (h_1, \dots, h_m)^T$  – вектор скрытых переменных
- $\Theta = (W, \mathbf{b}_v, \mathbf{b}_h)$  – параметры модели
- $W = (w_{ij})^{n \times m}$  – матрица бинарных потерниалов
- $\mathbf{b}_v, \mathbf{b}_h$  – векторы унарных потенциалов для видимого и скрытого слоёв соответственно
- $E(\mathbf{v}, \mathbf{h}, \Theta)$  – функционал энергии
- $Z(\Theta)$  – нормировочная константа

# Задача вывода

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Пусть дана обученная модель  $\Theta$  и входной вектор  $\mathbf{x}$ . Под задачей вывода понимаем:

- Оценка  $P(\mathbf{x}|\Theta)$ .
- Генерация объектов  $\mathbf{v}$  из распределения  $P(\mathbf{v}|\Theta)$ .
- Объект  $\mathbf{x}$  задан только частично, т.е. известны  $x_i, i \in I$ .  
Задача: оценить, либо генерировать из  $P(\mathbf{x}_{j \notin I} | \mathbf{x}_{i \in I}, \Theta)$ .

# Вывод в RBM

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

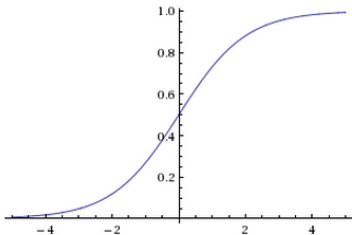
Закключение

Векторные формулы для получения условных вероятностей переменных слоёв  $\mathbf{v}$  и  $\mathbf{h}$ :

$$(P(v_i = 1 | \mathbf{h}, \Theta))_i^n = \sigma(-\mathbf{b}_v - W\mathbf{h})$$

$$(P(h_j = 1 | \mathbf{v}, \Theta))_j^m = \sigma(-\mathbf{b}_h - W^T \mathbf{v})$$

$$\sigma(x) = \frac{1}{1 + \exp(-x)}$$



# Вывод в RBM

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

$$\log P(\mathbf{x}|\Theta) = \log \sum_h P(\mathbf{x}, \mathbf{h}|\Theta)$$

$$-\log Z(\Theta) - \mathbf{b}_v^T \mathbf{x} - \sum_{i=1}^{nH} \left( \log \left( 1 + \exp \left( \left[ -\mathbf{x}^T W - \mathbf{h}^T \right]_{(i)} \right) \right) \right)$$

$$-\log Z(\Theta) - \log FE(\mathbf{x}, \Theta)$$

# Генерация выборки в RBM

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Для генерации данных можем использовать МСМС. Схема Гиббса в данном случае представляется следующим образом:

- $\mathbf{v}_0$  – начальная точка
- В силу условной независимости переменных  $\mathbf{h}$ , вектор  $\mathbf{h}_k$  генерируется из  $\mathbf{v}_k$  параллельно
- Затем генерируем  $\mathbf{v}_{k+1}$  из  $\mathbf{h}_k$ .  
Проблема: распределение  $P(\mathbf{v}, \mathbf{h}|\Theta)$  мультимодально, процесс часто застревает в локальных оптимумах, "перемешивание" цепей довольно слабо.



# Parallel Tempering

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Альтернатива схеме Гиббса, Parallel Tempering. Рассмотрим инверсные температуры  $0 = T_1 < \dots < T_N = 1$  и соответствующие им модели  $\Theta_k = \{W \cdot T_k, \mathbf{b}_v \cdot T_k, \mathbf{b}_h \cdot T_k\}$ . Строим  $N$  цепей Гиббса,  $k$ -я цепь соответствует модели  $\Theta_k$ . На каждом шаге получаем набор  $(\mathbf{x}_1, \dots, \mathbf{x}_N)$ . По циклу  $k = 2 \dots N$ : меняем  $\mathbf{x}_k$  и  $\mathbf{x}_{k-1}$  с вероятностью  $P_{swap}(\mathbf{x}_k, \mathbf{x}_{k-1})$

$$P_{swap}(\mathbf{x}_{T_1}, \mathbf{x}_{T_2}) = \min \left( 1, \frac{P_{T_1}(\mathbf{x}_{T_2})P_{T_2}(\mathbf{x}_{T_1})}{P_{T_1}(\mathbf{x}_{T_1})P_{T_2}(\mathbf{x}_{T_2})} \right)$$

Parallel Tempering is Efficient for Learning Restricted Boltzmann Machines. KyungHyun Cho, Tapani Raiko, Alexander Ilin

# Вариационный вывод

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Рассмотрим MF-аппроксимацию.

$$Q(\mathbf{v}, \mathbf{h} | \mu, \lambda) = \prod_{i=1}^{nV} \mu_i^{v_i} (1 - \mu_i)^{1-v_i} \prod_{i=1}^{nV} \lambda_i^{v_i} (1 - \lambda_i)^{1-v_i}$$

Тогда EM алгоритм:

$$\lambda := \sigma(-\mathbf{b}_h - W\mu)$$

$$\mu := \sigma(-\mathbf{b}_v - W^T \lambda)$$

# Задача обучения

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Пусть задана выборка одинаково распределённых независимых случайных величин  $X^I = (\mathbf{x}_1, \dots, \mathbf{x}_I)$ . Хотим найти параметры модели  $\Theta$  так, чтобы модель описывала входные данные наилучшим образом:

$$P(X^I, \Theta) = \prod_{i=1}^I P(\mathbf{v} = \mathbf{x}_i, \Theta) \rightarrow \max_{\Theta}$$

# Задача обучения

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Градиент функционала оптимизации:

$$\nabla_{\Theta} \log \sum_{i=1}^I P(\mathbf{v} = \mathbf{x}_i) = \nabla_{\Theta}^+ - \nabla_{\Theta}^- + \nabla \log P(\Theta)$$

$$\nabla_{\Theta}^+ = \sum_{i=1}^I \sum_{\mathbf{h}} P(\mathbf{h} | \mathbf{v} = \mathbf{x}_i, \Theta) \nabla_{\Theta} (-E(\mathbf{x}_i, \mathbf{h}, \Theta))$$

$$\nabla_{\Theta}^- = \sum_{\mathbf{v}, \mathbf{h}} P(\mathbf{v}, \mathbf{h} | \Theta) \nabla_{\Theta} (-E(\mathbf{v}, \mathbf{h}, \Theta))$$

Здесь  $\nabla_{\Theta}^+$  вычисляется аналитически, но оценка  $\nabla_{\Theta}^-$  – NP-полная задача.

# Задача обучения

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

$$\begin{aligned}\nabla_{W,i}^+ &= \mathbb{E}_{(\mathbf{v}, \mathbf{h}) \sim P(\mathbf{h}|\mathbf{v}, \Theta)} \delta(\mathbf{v} - \mathbf{x}_i) \nabla_W (-E(\mathbf{v}, \mathbf{h}, \Theta)) = \\ & \mathbb{E}_{data=\mathbf{x}_i} \mathbf{v} \mathbf{h}^T = \mathbf{x}_i \mathbb{E}_{data=\mathbf{x}_i} \mathbf{h}^T = \mathbf{x}_i \sigma(-\mathbf{b}_v - \mathbf{x}_i W)^T\end{aligned}$$

$$\nabla_{\mathbf{b}_v, j}^+ = \mathbb{E}_{data=\mathbf{x}_i} \mathbf{v} = \mathbf{x}_i$$

$$\nabla_{\mathbf{b}_h, j}^+ = \mathbb{E}_{data=\mathbf{x}_i} \mathbf{h} = \sigma(-\mathbf{b}_h - W^T \mathbf{x}_i)$$

# Contrastive Divergence

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

До сих пор целью оптимизации был функционал  
 $\log P(X^l, \Theta) = \log P(X^l | \Theta) + \log P(\Theta)$ .

Задача  $\log P(X^l | \Theta) \rightarrow \max_{\Theta}$  эквивалентна минимизации  
дивергенции Кульбака-Лейблера

$KL(Q_{data}^0(\mathbf{v}, \mathbf{h}) || Q_{\Theta}^{\infty}(\mathbf{v}, \mathbf{h})) \rightarrow \min_{\Theta}$ .

$$Q_{data}^0(\mathbf{v}, \mathbf{h}) = \frac{1}{l} \sum_{i=1}^l P(\mathbf{h} | \mathbf{v}, \Theta) \delta(\mathbf{v} - \mathbf{x}_i)$$

$$Q_{\Theta}^{\infty}(\mathbf{v}, \mathbf{h}) = P(\mathbf{v}, \mathbf{h} | \Theta)$$

# Contrastive Divergence

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Hinton в статье "Training Products of Experts by Minimizing Contrastive Divergence" предложил эвристику: оценивать статистику модели приближенно из распределения, получающегося из данных с помощью схемы Гиббса.

$$Q(\mathbf{v}, \mathbf{h})^0 \rightarrow Q(\mathbf{v}, \mathbf{h})^1 \rightarrow \dots$$

Вместо исходной задачи

$$KL(Q(\mathbf{v}, \mathbf{h})^0 \| Q(\mathbf{v}, \mathbf{h})^\infty) \rightarrow \min_{\Theta}$$

на каждом шаге градиентного спуска решаем

$$CD_N = KL(Q^0 \| Q^\infty) - KL(Q^N \| Q^\infty) \rightarrow \min_{\Theta}$$

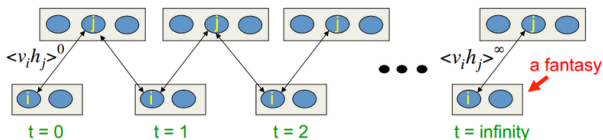
# Contrastive Divergence

$$\nabla_{\Theta} \log CD_N = (\nabla_{\Theta, i}^+ - \nabla_{\Theta}^-) - (\nabla_{\Theta, i}^{Q^N} - \nabla_{\Theta}^-) =$$

$$\nabla_{\Theta, i}^+ - \nabla_{\Theta, i}^{Q^N} = \mathbb{E}_{Q^0} \nabla(-E) - \mathbb{E}_{Q^N} \nabla(-E)$$

$$Q^N(\mathbf{v}, \mathbf{h}) = \frac{1}{I} \sum_{i=1}^M P(\mathbf{h}|\mathbf{v}, \Theta) \delta(\mathbf{v} - \mathbf{v}_i^{(N)})$$

Здесь  $Q^N$  получается из  $Q^0$  после  $N$  итераций схемы Гиббса.





# Persistent Contrastive Divergence

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Задача обучения RBM обычно решается градиентными методами. В процессе спуска, по мере приближения к точке оптимума, модель изменяется слабо. Это значит, что распределение  $Q_{t-1}^N$  на предыдущем шаге близко к  $Q_t^N$  на текущем (при условии фиксированного randseed). В алгоритме Persistent Contrastive Divergence предлагается начинать цепь не с  $Q^0$ , а с  $Q_{t-1}^N$ , так как  $Q_{t-1}^N$  находится ближе к  $Q_t^\infty$ , чем  $Q^0$  рядом с точкой оптимума.

# Метрики оценки качества

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение

Модели, обученные с при одинаковом априорном распределении с одинаковыми значениями гиперпараметров можно сравнивать по значению логарифма правдоподобия. Но! Это некорректно в противном случае. Метрика MSER (MSE on Reconstruction):

$$\mathbf{h}_{rec} \sim P(\mathbf{h}|\mathbf{v} = \mathbf{x}_{orig}, \Theta)$$

$$\mathbf{x}_{rec} \sim P(\mathbf{v}|\mathbf{h} = \mathbf{h}_{rec}, \Theta)$$

$$MSER(\mathbf{x}_{orig}, \Theta) = \|\mathbf{x}_{orig} - \mathbf{x}_{rec}\|$$

Смысл – насколько хорошо модель восстанавливает данные при передаче информации о исходном видимом векторе на скрытый слой.

# Архитектура

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

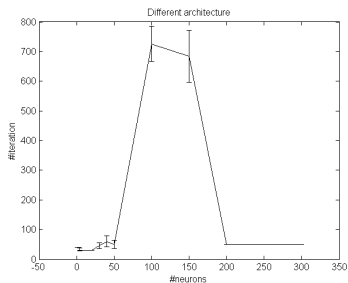
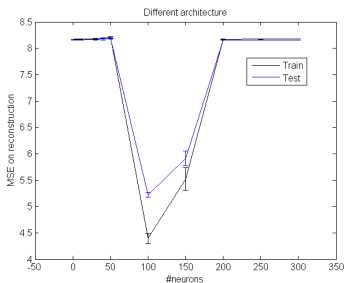
Задача  
обучения

Экспери-  
менты

**Архитектура**

Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Закключение



Uniform prior

# Архитектура

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограниченной  
машины  
Больцмана

Задача  
вывода

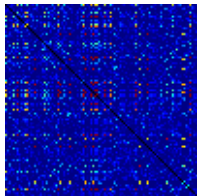
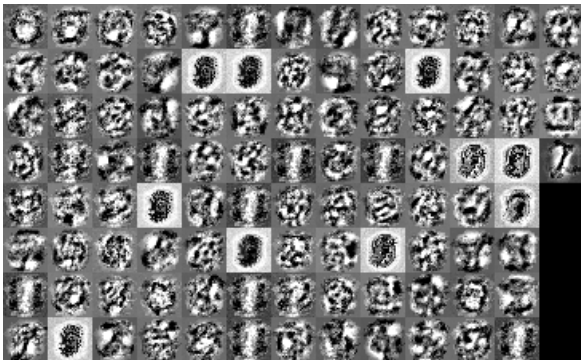
Задача  
обучения

Экспери-  
менты

**Архитектура**

Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение



100 нейронов

# Архитектура

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

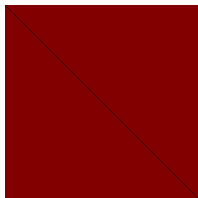
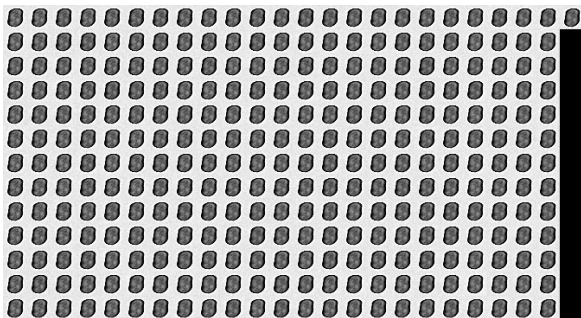
Задача  
обучения

Экспери-  
менты

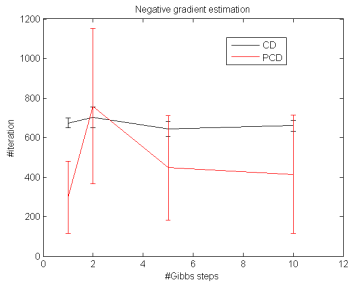
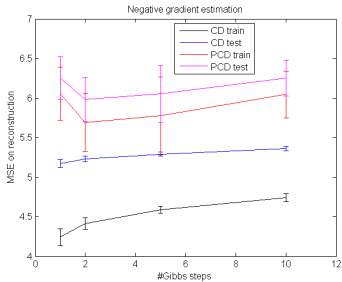
**Архитектура**

Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение



300 нейронов



# Нормальное распределение

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура

Оценка  
градиента

**Регуляри-  
заторы**

Вывод

Заключение

Рассмотрим регуляризацию с помощью нормального априорного распределения (L2-регуляризация)

$$W \sim \mathcal{N}(\Sigma = \alpha^{-1}I, \mu = 0)$$

$$\nabla_{\Theta} \log P(\Theta) = -\alpha W$$

# Нормальное распределение, 50 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

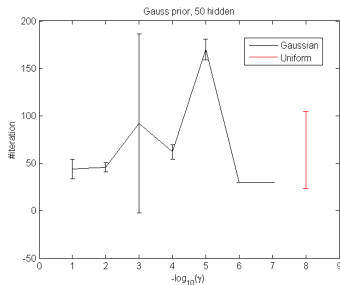
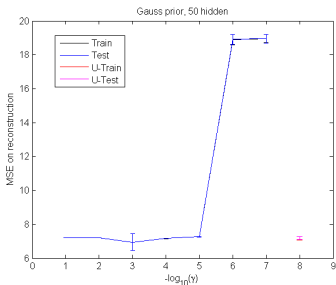
Архитектура

Оценка  
градиента

Регуляри-  
заторы

Вывод

Закключение





# Нормальное распределение, 50 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

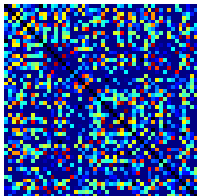
Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
**Регуляри-  
заторы**  
Вывод

Заключение



$$\gamma = 10^{-3}$$

# Нормальное распределение, 100 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

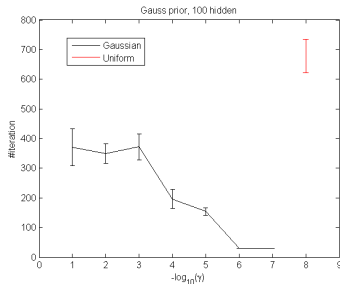
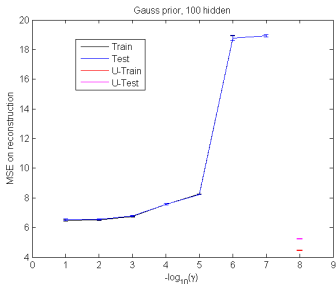
Архитектура

Оценка  
градиента

**Регуляри-  
заторы**

Вывод

Заключение



# Нормальное распределение, 300 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

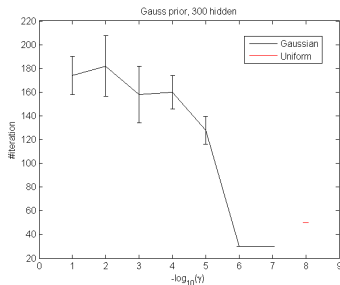
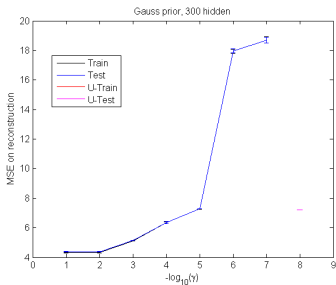
Архитектура

Оценка  
градиента

**Регуляри-  
заторы**

Вывод

Закключение



# Нормальное распределение, 300 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

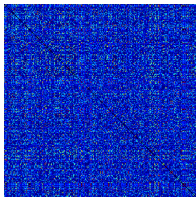
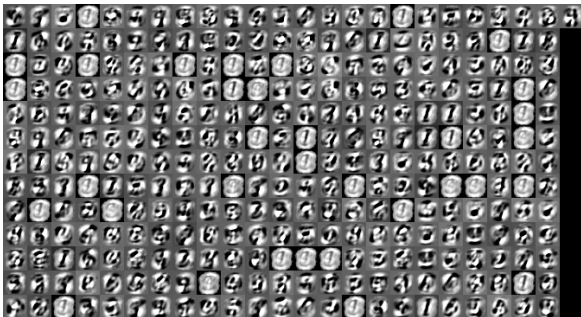
Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
**Регуляри-  
заторы**  
Вывод

Заключение



Лучшая модель

# Нормальное распределение, 300 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

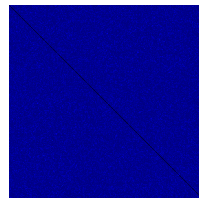
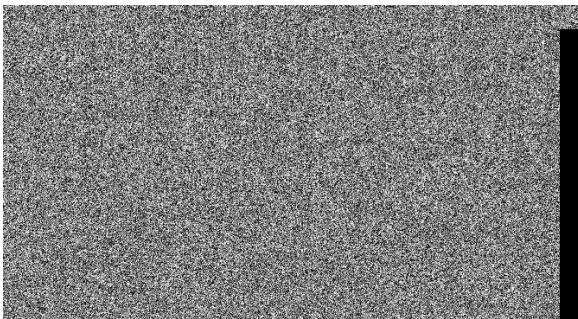
Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
**Регуляри-  
заторы**  
Вывод

Заключение



$$\gamma = 10^{-6}$$

# Распределение Лапласа

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента

**Регуляри-  
заторы**

Вывод

Заключение

Рассмотрим регуляризацию с помощью априорного распределения Лапласа (L1-регуляризация)

$$W_{i,j} \sim \text{Laplace}(\alpha^{-1}, \mu = 0)$$

$$\nabla_{\Theta} \log P(\Theta) = -\alpha \text{sign}(W)$$

# Распределение Лапласа, 50 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

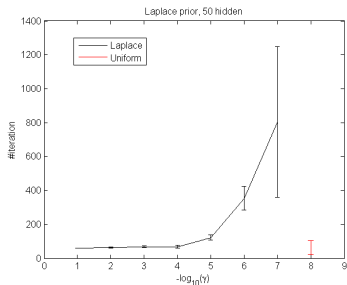
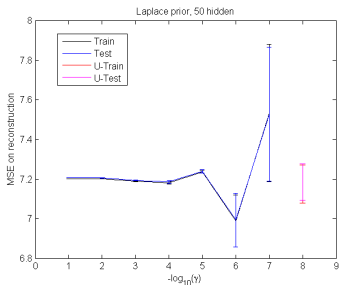
Архитектура

Оценка  
градиента

**Регуляри-  
заторы**

Вывод

Закключение



# Распределение Лапласа, 100 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

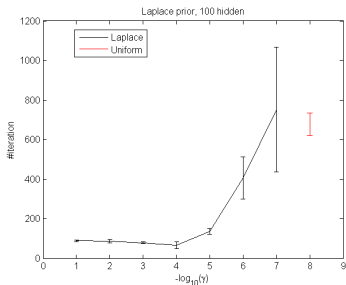
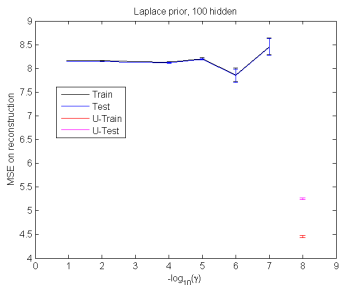
Архитектура

Оценка  
градиента

**Регуляри-  
заторы**

Вывод

Закключение





# Распределение Лапласа, 300 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

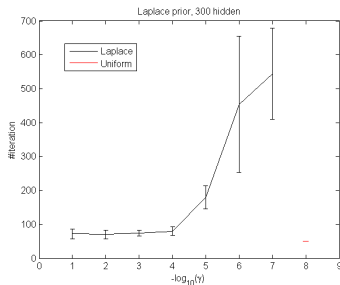
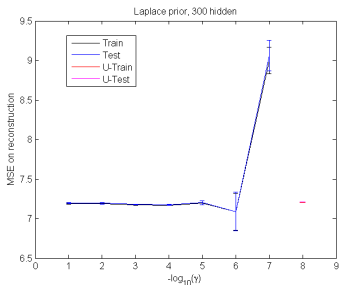
Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
**Регуляри-  
заторы**  
Вывод

Заключение



# Распределение Лапласа, 300 нейронов

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

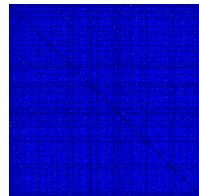
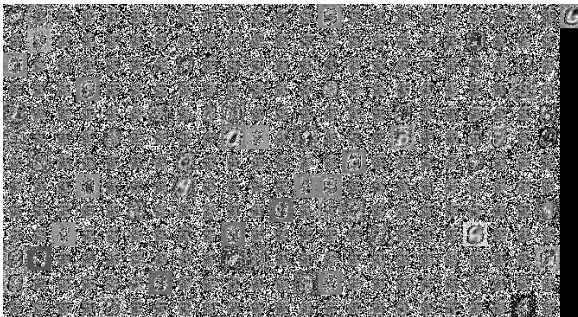
Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение



Лучшая модель,  $\gamma = 10^{-6}$

# Восстановление изображений

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
**Вывод**

Заключение

Попробуем восстановить цифру '6', порезанную пополам



Рис. : Результат после вывода схемой Гиббса

# Генерация новых данных

Ограниченные  
Машины  
Больцмана

Хомутов  
Н.Ю.

Модель  
ограничен-  
ной машины  
Больцмана

Задача  
вывода

Задача  
обучения

Экспери-  
менты

Архитектура  
Оценка  
градиента  
Регуляри-  
заторы  
Вывод

Заключение



Рис. : Новые изображения, сгенерированные из белого шума

+ видео вывода для схем PT, Gibbs, variational

# Спасибо за внимание!