Hidden Makov Model

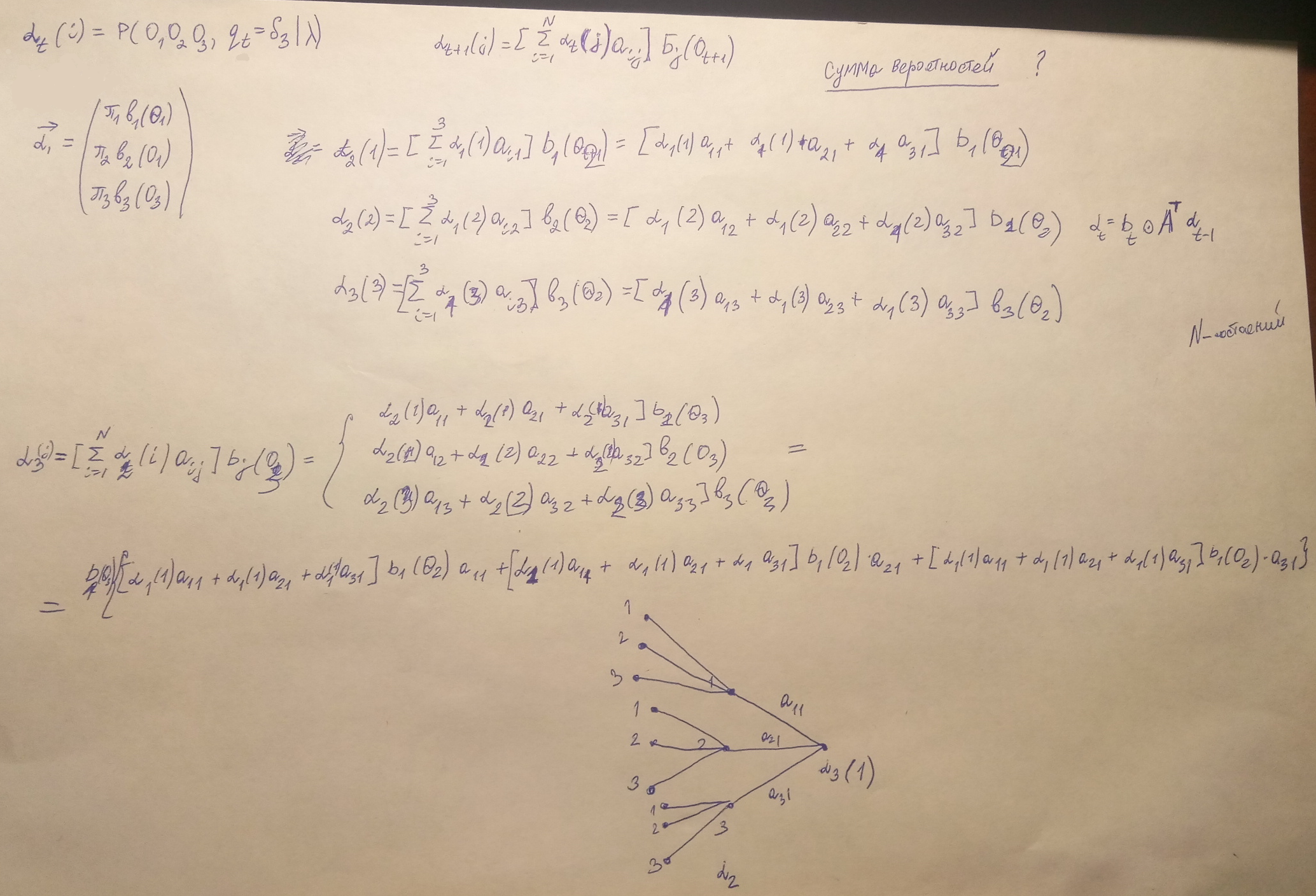
Источники:

1. Kevin P.Murphy Machine Learning A Probabilisti. [589-628]

Алгоритмы обучения

Выбор количества скрытых состояний? [Murhy 621]

**Forward algorithm**



Baum Welch

Использует EM [Murhy 348]

Viterbi

Присваивает жесткие метки каждому наблюдению

Pomegranate

Тип входных данных для обучения:

seq=numpy.array(list('CGACTACTGACTACTCGCCGACGCGACTGCCGTCTATACTGCGCATACGGC'))

shape = (51,)

HiddenMarkovModel(GraphModel):

dense\_transition\_matrix() - возвращает матрицу переходов

from\_samples():

#

n\_components — количество состояний в модели

#

проверка n\_components != мощность словаря диск. наблюдений

Задачи 24.12.18

1. Смоделировать нормальные последовательности (Определить вид).

Типы:

Последовательная. (abcd)

aaabbcde

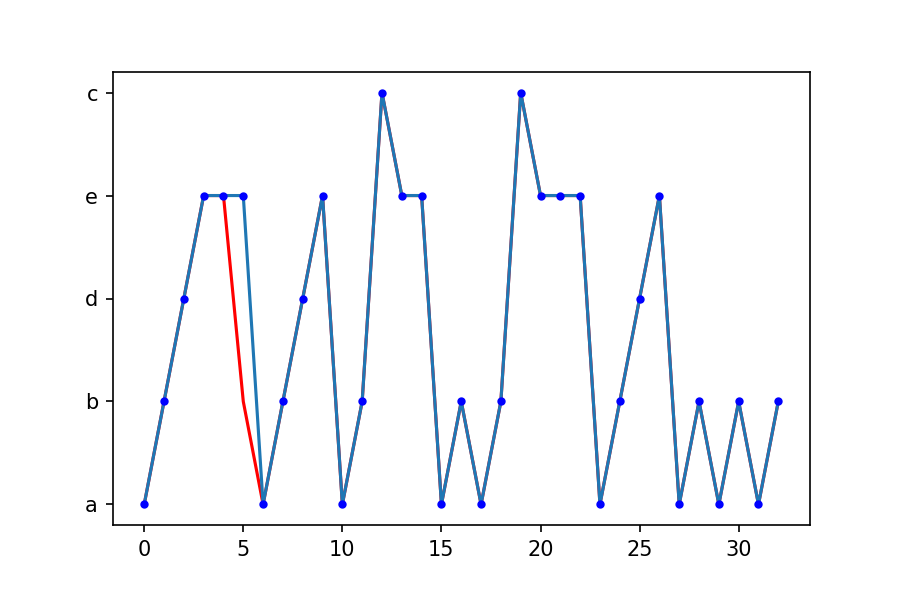
2. Получить аномальные.

3. Получить значение логарифма правдоподобия.

4. Построить графики сделать выводы.

Эксперимент №1

Дискретный сигнал, с невозможный переходом для марковской цепи ( MarkovChain)



['a', 'a', '0.0']

['a', 'b', '1.0']

['a', 'd', '0.0']

['a', 'c', '0.0']

['a', 'e', '0.0']

['b', 'a', '0.375']

['b', 'b', '0.0']

['b', 'd', '0.375']

['b', 'c', '0.25']

['b', 'e', '0.0']

['d', 'a', '0.0']

['d', 'b', '0.0']

['d', 'd', '0.0']

['d', 'c', '0.0']

['d', 'e', '1.0']

['c', 'a', '0.0']

['c', 'b', '0.0']

['c', 'd', '0.0']

['c', 'c', '0.0']

['c', 'e', '1.0']

['e', 'a', '0.5'] Нормальный -15.589036045909591, Аномальный -inf

['e', 'b', '0.0']

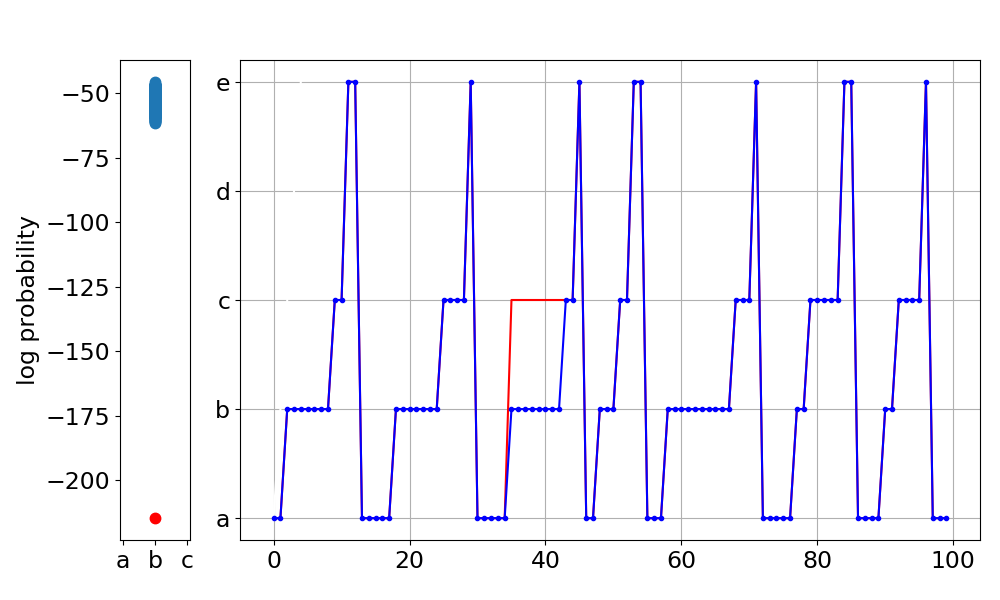
['e', 'd', '0.0']

['e', 'c', '0.0']

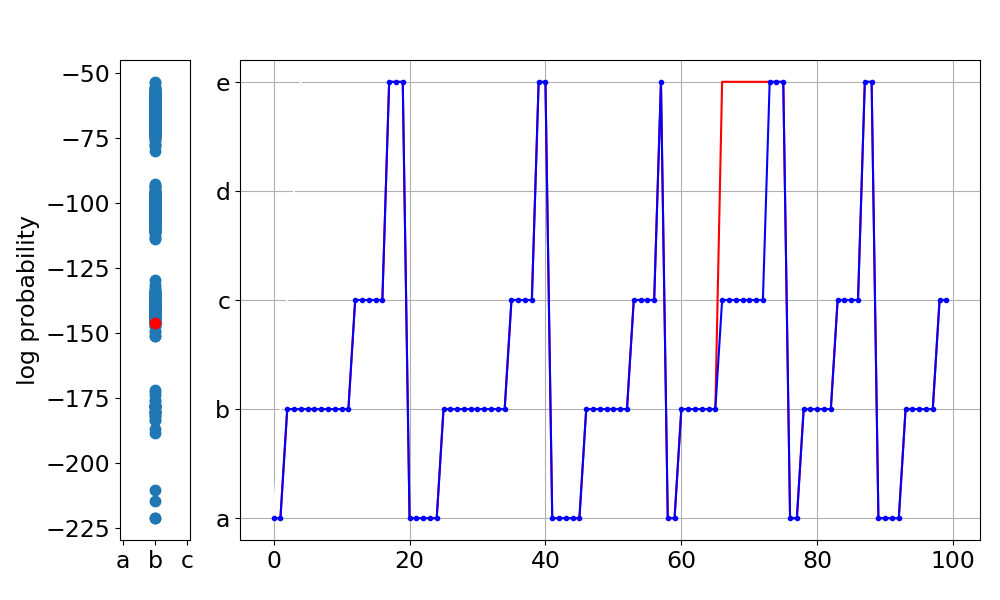
['e', 'e', '0.5']

Эксперимент №2

Дискретный сигнал, с внесением аномалии в последовательность для Скрытой марковской цепи (hmm).

**- HMM ( без меток ), 5 состояний, дискретное распределение.**

Значение логарифма вероятности ниже для последовательности с аномальной вставкой.



Однако возможна ситуация когда, значение лог. Вер. Для аномальной последовательности не ниже, чем для других последовательностей. Почему? (Так определились вероятности что вер. перехода из b в e велика.)

Эксперимент

1. Сгенерировали сигнал.

2. Построили модель.

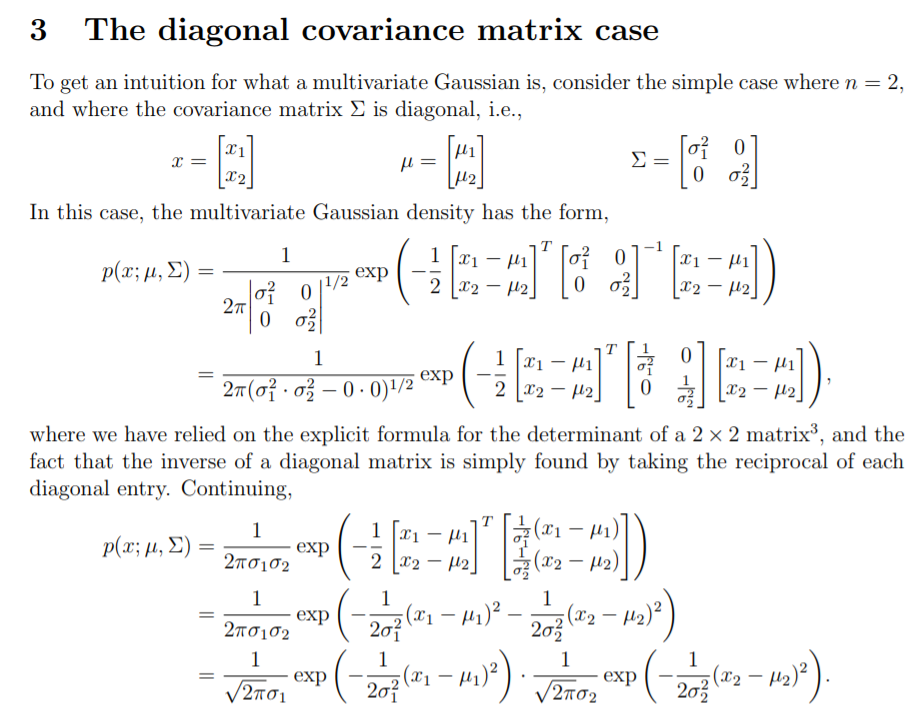
3. Подсовываем значения определяем состояния.

- Значения из распределения как на обучении

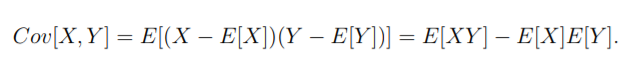
- Значения из другого распределения.

Распознавание изображений MNIST

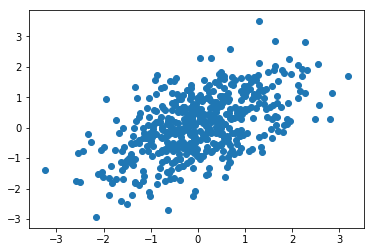
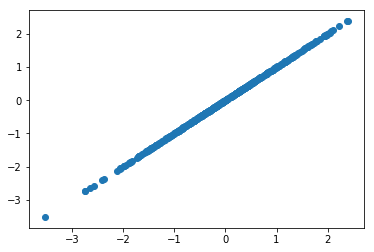
Многомерное гауссово распределение



Матрица ковариаций



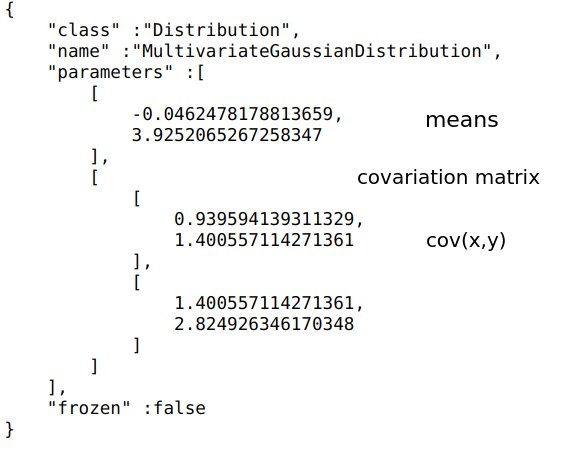
m = [0,0] cov = [[1,1], [1,1]] m = [0,0] cov = [[1,0.5, [0.5,1]]



Положительно определенность матрицы нарушается в данном случае если cov(x,y) >1

Pomegranate

MultivariateGaussianDistribution



09.01.

1. Научиться обучать gmm под сигнала и размечать состояния через него.

2. Функция plot scatter.

Случайно четыре признака.

16.01

1. Разобраться с тиками для построения последовательностей.

Решение: Сначала отрисовать алфавит: ax.plot(['a','b','c','d','f'],'w')

17.01

1. Вывод модели состояние.наблюдения

18.

Проверить работу алгоритма hmm c разметкой:

1. Сгенерировать последовательность.

2. Путь меток.

3. Построить графики

4. Вывести матрицу анализ

5. Ошибка.

06.02.18

1. Функция для генерации сигнала по меткам. ( Различная дисперсия и мат. Ожидание?).

Возможные графики для статьи

1. Множество сгенерированных дискретных сигналов.

2. Похожие сигналы на продолжительность нахождения в одном состоянии разная.

Для семинара

1. Формулы расчета вероятности отдельный график.

2. Для дискретных с разной продолжительностью в состоянии.

3. Для непрерывных с разной дисперсией.

4. Для многомерных.

Процедура:

1. Получить сигнал и метки скрытого состояния.

2. Обучить модель.

3. Получить аномальный сигнал.

4. Получить последовательность скрытых состояний.

5. Подсчитать вероятность. Если вероятность == 0, то аномалия.

get\_model( signal, label)

get\_anomal ( model, anomal\_signal)

**Разбор статьи**

**Multivariate time series anomaly detection: A framework of Hidden Markov Models**

1. z-score нормализация временного ряда.

2. HMM c 2-мя состояниями (Нормальное аномальное)

