Цель работы: алгоритм непрерывного структурного обучения динамических байесовских сетей

Задачи:

1. Сбор и анализ литературы

Введение

Прежде чем переходить к динамическим байесовским сетям и их изучению очень важно понимать, с чем же именно мы будем работать. Какова природа процесса? Какие инструменты существуют, чтобы их описывать? Каковы их ограничения, какие преимущества?

В области ML существует очень много инструментов, которые позволяют работать со стационарными процессами. Под стационарностью мы будем понимать неизменность во времени[]. Все эти инструменты хорошо описаны и имеют строгую математическую базу[]. Дело же с нестационарными, то есть динамическими процессами, обстоит иначе.

Для описания нестационарных процессов традиционно используют аппараты либо теории стохастических процессов или же аппарат динамических систем[]; главное различие в этих подходах – это отсутствие детерминизма и его наличие, соответственно.

Стохастические процессы

Стохастический (случайные) процесс представляет собой множество случайных величин, которые каким-либо образом меняются во времени. Для каждой случайной величины существует одно и только одно t T.

Теория случайных процессов является расширением математической статистике, в котором случайные величины зависят не только от элементарного исхода, но и еще от параметра – времени.

Рассмотрим переход от простой случайной величины к темпоральной и какие понятия такой переход порождает.

1. *ξ*(*ω*) переходит в *ξ*(*ω, t*), где *ω –* элементарный исход*, ξ* – случайная величина, *t* – динамичская переменная. Если зафиксировать время *t0*, то будет снова случайная величина.
2. Для каждого t T существует свое распределение. В классической теории рассматривается только те случаи, когда весь процесс задан одним классом распределений и для каждого *t* свой набор параметров.
3. Если из каждого распределения взять по выборке размером 1, то получится временной ряд, который называется реализацией процесса. Множество всех возможных реализаций называется ансамблем.

В зависимости от T процессы делятся: <…>

Ограничения стохастических процессов

Все случайные величины одного процесса должны принадлежать одному пространству состояний. Что само по себе уже является серьезным ограничением, поскольку в реальности, как правило, имеется динамическая ось (например, время, глубина, длина волны) и множество реализаций разных процессов, которые связаны между собой. Моделирование процесса как случайного предполагает существование совместного распределения, которое не является факторизованным, то есть, например, в случае гауссовского процесса вместо рассмотрения одного среднего и одной дисперсии рассматривается матрица. В общем случае, вопрос о независимостях возникает лишь постфактум, когда совместное распредление известно, за исключением марковского процесса, в котором постулируется попарная обусловленность между Xt и Xt+1 для любого t T.

Несмотря на ограничение выше, такая теория оказывается приближенной к ограниченной (но не упрощенной) реальности. Например, можно рассмотреть количество частиц от распада радиоактивного вещества, которое регистрирует датчик. Вероятность столкновения с датчиком крайне мала, но частиц много, поэтому из закона редких событий делается предположение о пуассоновском распределении частиц на интервале времени. Весь процесс в таком случае будет пуассоновским.

Другая крайне важная проблема кроется в связи между случайностью и самой реальностью. Понимая динамику исключительно как ансамбль возможных состояний, мы попадаем в ловушку статистической физики.

Рассмотрим эту проблему с двух крайностей.

1. Концентрация. Когда химик говорит, что раствор карбоната натрия имеет концентрацию 1 моль/литр, это означает, что каким бы образом он ни отобрал 0.001 л (1 мл) из этого раствора, его концентрация будет пропорциональна исходному раствору. Количество частиц настолько велико, что флуктуации концентрации будут незначительны. Однако же представим себе концентрацию очень маленькую, например 10 молекул карбонат натрия на 6 \* 1023 молекул воды (10 молекул / 1 моль воды). Рассуждения по аналогии выше будут уже неверны, ведь флуктуации концентрации будут столь велики, что среднее по ансамблю за тот же интервал времени будет уже не равномерным. Естественно (как любят делать химики), обычно такой маленькой концентрацией просто пренебрегают и получают нужный результат из расчетов, однако наша же задача показать, что среднее по ансамблю не всегда будет работать.
2. Сложное поведение. Допустим, что количество частиц теперь достаточно, чтобы брать среднее по ансамблю, и случай, кажется, хороший. Но почему-то в растворе частицы перестают вести себя статистически, как материальные точки. Например, химик работает с раствором какой-то длинной молекулы (но не с полимером). И неожиданно раствор вместо того, чтобы проявлять свойства статистические, начинает проявлять свойства динамические: при нагревании возникают оптические свойства, которые никак не выводятся из свойств отдельной молекулы. Для нас это означает, что рассматривая только стохастический процесс, мы рискуем упустить очень важный элемент – динамику. Иными словам, нам следует задать вопрос: каково взаимодействие между ансамблями? Действительно ли набора случайных переменных достаточно, чтобы описать все возможные ансамбли?

[НУЖНО РАССМОТРЕТЬ ОБУЧЕНИЕ РАСПРЕДЕЛЕНИЯ]

Динамические системы

Динамическая система – это способ анализа системы, при котором полагается, что каждое новое состояние системы зависит от 2 факторов: предыдущего значения, которое накапливает в себе динамику, а также от времени и вектора параметров. В связи с этим, существует два математических объекта, которые описывают их: дифференциальные уравнения и разностные уравнения.

Само определение уже порождает сложности, потому что в классических моделях обычно работают с аналитическими объектами, в динамических системах же приходится работать с функциями на каком-либо интервале (для дифференциального уравнения – на бесконечно малом дифференциале, для разностного – на конечном отрезке динамической оси). Однако же, если удается найти его, то у него проявляются удивительные свойства, которые сходятся с реальностью[].

[Примеры динамической системы]

[Как динамическая система может породить временной ряд?]