Модель типа Лотки–Вольтерры с мутациями и порождающие соревновательные сети

Обобщение модели теории машинного обучения, модели порождающей соревновательной сети (GAN), на случай популяции GAN.

Хищник и жертва в модели Лотки–Вольтерры описывают генератор и дискриминатор в модели GAN, адаптация описывает обучение.

Модель описывает контроль переобучения для GAN с использованием подхода алгоритмической устойчивости (популяция GAN будет стремиться занять широкие максимумы функции правдоподобия).

Теория игр. Игра есть функция нескольких переменных, называемых игроками, игра принимает значения, равные набору вещественных чисел, по одному для каждого игрока

$$(v_1, \dots, v_n)(s_1, \dots, s_n),$$

v_i – выигрыш i-го игрока;
s_i — стратегия i-го игрока.

Пример: два игрока, игра с нулевой суммой
(сумма выигрышей игроков равна нулю).

Смешанная стратегия для i-го игрока: распределение вероятностей $p_i(s^i_j)$ для стратегий этого игрока.
Выигрыш i-го игрока для набора смешанных стратегий

$$\langle v_i \rangle = \sum_{j_1,\ldots,j_n} p_1(s^1_{j_1}) \ldots p_n(s^n_{j_n}) v_i(s^1_{j_1}, \ldots, s^n_{j_n})$$

есть усреднение выигрыша i-го игрока по смешанным стратегиям для всех игроков. Здесь индекс j_k пробегает значения возможных стратегий s_k для k-го игрока:
$s_k \in \{s^k_1, \ldots, s^k_{m_k}\}$.

Равновесие Нэша: набор стратегий, в котором ни один участник не может увеличить выигрыш, изменив свою стратегию, если другие участники своих стратегий не меняют.

Максимин: наибольший выигрыш, который данный игрок может получить, не зная действий других игроков

$$\underline{v_i} = \max_{s_i} \min_{s_{-i}} v_i(s_i, s_{-i})$$

s_i – стратегия i-го игрока;
s_{-i} – стратегии остальных игроков;
v_i – выигрыш i-го игрока.

Минимакс: наименьший выигрыш, который другие игроки могут заставить получить, не зная действия данного игрока

$$\overline{v_i} = \min_{s_{-i}} \max_{s_i} v_i(s_i, s_{-i}).$$

Минимакс не меньше максимина

$$\overline{v_i} \geq \underline{v_i};$$

Минимакс для игр двух игроков с нулевой суммой совпадает с равновесием Нэша.

Модель Лотки–Вольтерры описывает популяционную динамику двух видов (хищник и жертва) системой двух ОДУ

$$\frac{dx}{dt} = \alpha x - \beta xy;$$

$$\frac{dy}{dt} = -\gamma y + \delta xy;$$

x – численность жертвы;

y – численность хищника.

Все константы положительны; нелинейный член описывает взаимодействие хищника и жертвы.

Для экологической ниши конечного объёма первое уравнение меняется как

$$\frac{dx}{dt} = \alpha(x - x^2/N) - \beta xy.$$

Модель Лотки–Вольтерры с мутациями:
жертвы x_i, хищники y_m, уравнения

$$\frac{dx_i}{dt} = \sum_j A_{ij} x_j - \sum_n B_{in} x_i y_n;$$

$$\frac{dy_m}{dt} = \sum_n C_{mn} y_n + \sum_j B_{jm} x_j y_m.$$

Матрица A – диагональная часть размножение жертв, внедиагональная мутации;
Матрица C – диагональная часть вымирание хищников, внедиагональная мутации;
Матрица B – взаимодействие хищника и жертвы.
Для ограниченной экологической ниши, аналогично первое уравнение примет вид

$$\frac{dx_i}{dt} = \sum_j A_{ij} x_j - \frac{1}{N} x_i \sum_{ij} A_{ij} x_j - \sum_n B_{in} x_i y_n.$$

Информационная геометрия

Статистическое многообразие — многообразие параметров параметрического распределения вероятности, или многообразие, точки которого являются распределениями вероятности на X.

Расстояние Кульбака–Лейблера (несимметричное) между двумя распределениями вероятностей на X:

$$D(p|q) = \int_X p(x) \log \frac{p(x)}{q(x)} dx.$$

Метрика Фишера–Рао на статистическом многообразии

$$g_{ij}(\theta) = \int_X \frac{\partial \log p(x,\theta)}{\partial \theta_i} \frac{\partial \log p(x,\theta)}{\partial \theta_j} p(x,\theta) dx.$$

Разложение расстояния Кульбака–Лейблера для малых $\Delta\theta = \theta - \theta_0$ через метрику Фишера:

$$D(p(\theta_0)|p(\theta)) = \frac{1}{2}\sum_{ij} g_{ij}(\theta_0)\Delta\theta^i\Delta\theta^j.$$

Оператор Лапласа–Бельтрами на статистическом многообразии с параметром θ, g_{ij} есть метрика Фишера

$$\Delta_\theta = \frac{1}{\sqrt{g}}\sum_i \frac{\partial}{\partial\theta_i}\left(\sqrt{g}\sum_j g^{ij}\frac{\partial}{\partial\theta_j}\right).$$

Метод максимального правдоподобия

Параметрическое семейство распределений вероятности $p(x, \theta)$, $x \in X$. Выборка независимых испытаний $\{x_i\}$ в X из некоторого распределения вероятности q. Функция правдоподобия

$$L(\{x_i\}, \theta) = \prod_{i=1}^{n} p(x_i, \theta).$$

Оценка максимального правдоподобия θ

$$\theta_0 = \arg\max L(\{x_i\}, \theta) =$$

$$= \arg\max \frac{1}{n} \log L(\{x_i\}, \theta) = \arg\max \frac{1}{n} \sum_{i=1}^{n} \log p(x_i, \theta)$$

принимает вид эмпирической суммы; распределение $p(x, \theta_0)$, отвечающее оценке максимального правдоподобия, приближает неизвестное распределение $q(x)$.

Обучение — нахождение гипотезы (параметра θ) максимального правдоподобия (минимального риска).

Переобучение есть отсутствие способности к обобщению — подгонка $p(x, \theta)$ к обучающей выборке $\{x_i\}$ (высокое правдоподобие), плохие результаты (низкое правдоподобие) на контрольной выборке $\{x_i'\}$ (некоторой другой выборке, порождённой распределением q).

Аналогия между обучением и дарвиновской эволюцией как между двумя задачами оптимизации — максимизация приспособленности как минимизация эмпирического риска (А.Тьюринг, 1950).

Популяционное обобщение модели обучения — по аналогии с переходом от отдельных систем в механике к ансамблям систем в статистической механике, и от эволюции организма к эволюции популяции в популяционной генетике.

Вместо обучения модели как оптимизации по параметру θ рассматривается популяция моделей обучения (с разными параметрами θ) и некоторая задача популяционной генетики (параметр θ есть аналог генома).
Популяционная динамика должна иметь вид сходимости популяции к окрестности решения задачи обучения (вместо отбора более приспособленного гена изучается изменение частоты генов в популяции).

Более сложные способы борьбы с переобучением.

Порождающие соревновательные сети
(generative adversarial networks, GANs):
X — пространство исследуемых объектов.
p_{data} есть обучающая выборка (распределение) на X.
Распределения:
Дискриминатор $D(x, \theta_d)$ с параметром θ_d;
Генератор $p_{\mathrm{gen}}(x, \theta_g)$ с параметром θ_g.

Задачи оптимизации: дискриминатор максимизирует близость к данным p_{data} и удалённость от генератора p_{gen}

$$V(\theta_d, \theta_g) = E_{x \sim p_{\mathrm{data}}} \log D(x, \theta_d) + E_{x \sim p_{\mathrm{gen}}} \log(1 - D(x, \theta_d)) =$$

$$= \int_X p_{\mathrm{data}}(x) \log D(x, \theta_d) dx + \int_X p_{\mathrm{gen}}(x, \theta_g) \log(1 - D(x, \theta_d)) dx. \tag{1}$$

Распределение данных есть выборка $\{x_i\}$, $i = 1, \ldots, n$, тогда функционал (1) принимает вид эмпирической суммы

$$V_{\mathrm{emp}}(\theta_d, \theta_g) = \frac{1}{n}\sum_{i=1}^{n} \log D(x_i, \theta_d) + \int_X p_{\mathrm{gen}}(x, \theta_g) \log(1 - D(x, \theta_d)) dx. \tag{2}$$

Генератор должен минимизировать величину (по θ_g)

$$E_{x \sim p_{\mathrm{gen}}} \log(1 - D(x, \theta_d)) = \int_X p_{\mathrm{gen}}(x, \theta_g) \log(1 - D(x, \theta_d)) dx.$$

Объединение этих двух задач дает следующую задачу из теории игр (проблема минимакса, равновесие Нэша): найти θ_d и θ_g, такие, что достигается минимакс

$$\min_{\theta_g} \max_{\theta_d} V_{\text{emp}}(\theta_d, \theta_g).$$

Распределение генератора $p_{\text{gen}}(x, \theta_g)$ (где θ_g есть решение задачи минимакса) есть решение задачи обучения для GAN.

Второй член в (2) (проигрыш генератора, добавка к эмпирической сумме) и проблема минимакса (максимизировать выигрыш дискриминатора, минимизировать проигрыш генератора) служат борьбе с переобучением в задаче нахождения распределения максимального правдоподобия.

GAN и эволюционная теория игр.
Обучающая выборка p_{data} есть распределение травы, дискриминатор D есть распределение травоядных, генератор p_{gen} есть распределение хищных.

Задачи оптимизации: D ищет траву и убегает от хищников, хищник p_{gen} ищет травоядных. Параметры θ_d и θ_g (гипотезы для дискриминатора и генератора) суть генотипы травоядных и хищных, вариация этих параметров означает наличие мутаций генотипов, задача оптимизации означает дарвиновскую эволюцию, адаптация есть сходимость дискриминатора и генератора к решению задачи минимакса для GAN.

Появление хищников вызывает ускорение эволюции.
Связь с контролем переобучения?

Популяционное обобщение модели GAN.

Модель Лотки–Вольтерры: дискриминатор есть жертва, которая пасётся на данных, а генератор есть хищник, который ловит жертву.

Статистическое многообразие есть пространство геномов, диффузия на этом пространстве описывает мутации.

X с параметром x есть пространство данных;

Y с параметром y — пространство гипотез дискриминатора;

Z с параметром z — пространство гипотез генератора.

Распределение данных на X;

популяция дискриминаторов $f(y,t)$ на Y;

популяция генераторов $g(z,t)$ на Z.

Динамика популяций дискриминатора и генератора описывается диффузией на Y и Z

$$\frac{\partial}{\partial t} f(y,t) = M_d \Delta_y f(y,t) + A(y) f(y,t) - f(y,t) \int_Z B(y,z) g(z,t) dz;$$

$$\frac{\partial}{\partial t} g(z,t) = M_g \Delta_z g(z,t) - C g(z,t) + g(z,t) \int_Y B(y,z) f(y,t) dy;$$

$M_d, M_g > 0$, член с $C > 0$ для генератора описывает вымирание хищника при отсутствии жертв.

Члены с функциями A и B берутся из задачи минимакса.

$A(y)$ описывает кормление дискриминатора на данных (функция правдоподобия, рост популяции дискриминатора при близости к данным)

$$A(y) = \exp\left(\int_X p_{\mathrm{data}}(x) \log D_y(x) dx \right) = \prod_{i=1}^{n} D_y(x_i).$$

Член $B(y,z)$ описывает хищничество генератора на дискриминаторе

$$B(y,z) = -\alpha \int_X p_{\mathrm{gen},z}(x) \log(1 - D_y(x)) dx, \quad \alpha > 0.$$

$B(y,z)$ уменьшается при удалении распределений дискриминатора от генератора. Узкие пики распределения дискриминатора в пространстве гипотез эффективнее поедаются более специализированным хищником, из-за чего дискриминатор и генератор занимают более широкие области на своих статистических многообразиях.

Модель типа Лотки–Вольтерры с мутациями для GAN, или LVM–GAN.

Сравните $B(y,z)$ с $D(p_{\mathrm{gen},z}|D_y)^{-1}$, где расстояние Кульбака–Лейблера

$$D(p|q) = \int_X p(x) \log \frac{p(x)}{q(x)} dx,$$

несимметричный аналог кулоновского взаимодействия. Для GAN генератор притягивается к дискриминатору, а дискриминатор от него отталкивается (взаимодействие несимметрично).

Аналог модели Эйгена в рамках информационной геометрии (нет хищников, ограниченная экологическая ниша)

$$\frac{\partial}{\partial t} f(y,t) = M_d \Delta_y f(y,t) + A(y) f(y,t) -$$

$$-\frac{1}{N} f(y,t) \int_Y \left(M_d \Delta_u + A(u) \right) f(u,t) du.$$

$A(y)$ может зависеть от времени, если данные зависят от времени. Урезанный вариант предыдущей модели.

Переобучение: высокое правдоподобие на обучающей выборке и низкое на контрольной (низкая способность к обобщению).

Контроль переобучения:

Конечная и малая VC-размерность (простая модель обучения)

V. N. Vapnik, The Nature of Statistical Learning Theory, Springer, 1998.

Плоские минимумы функционала риска

S.Hochreiter, J.Schmidhuber, Flat Minima, Neural Computation 9, 1–42 (1997).

Алгоритмическая устойчивость (устойчивость решения алгоритма обучения к возмущениям обучающей выборки)

O. Bousquet, A. Elisseeff, Stability and Generalization, Journal of Machine Learning Research, 2, 499–526 (2002).

S. Kutin, P. Niyogi, Almost-everywhere algorithmic stability and generalization error, in: Proc. of 18-th Conference on Uncertainty in Artificial Intelligence (UAI 2002), 2002, pp. 275–282. arXiv:1301.0579 (2002).

T. Poggio, R. Rifkin, S. Mukherjee, P. Niyogi, General conditions for predictivity in learning theory, Nature, 428, 419–422 (2004).

Контроль переобучения для **модели типа Эйгена**: для узких пиков популяции на переобученных гипотезах мутации на близлежащие гипотезы (где приспособленность низка) ведут к снижению пиков и уменьшению переобучения.
Возможная зависимость $A(y)$ (данных) от времени усиливает этот эффект — пик приспособленности может сдвинуться относительно пика популяции, что для узких пиков снижает приспособленность (это в точности эффект отсутствия алгоритмической устойчивости обучения).
Подобный механизм обсуждался Р.Фишером (книга "Генетическая теория естественного отбора", 1930).
Слишком высокая скорость мутаций ведёт к размыванию всех пиков, даже отвечающих правильной гипотезе (недообучение).
Слишком высокая скорость мутаций в рамках модели Эйгена обсуждается как катастрофа ошибок и выражается в падении эффективности очищающего отбора в популяционной генетике.

Контроль переобучения для **модели типа Лотки–Вольтерры с мутациями**: 1) то же, что для модели типа Эйгена;
2) Попадание хищника (генератора) на узкий пик популяции жертвы (дискриминатора) ведёт к размножению хищника и выеданию пика. $B(y,z)$ снижается при удалении распределения генератора от распределения дискриминатора, то есть широкие пики распределения дискриминатора в пространстве гипотез медленнее выедают, более специализированый хищник эффективнее охотится. Это заставляет дискриминатор занимать более широкие области в пространстве гипотез, близкие к обучающей выборке, генератор следует за дискриминатором (эффект несимметричного отталкивания дискриминатора от генератора).
Подавление узких пиков популяций дискриминатора и генератора приводит к снижению переобучения. Модель типа Лотки–Вольтерры с мутациями обеспечивает более эффективный контроль переобучения, чем модель типа Эйгена; это также объясняет ускорение биологической эволюции при наличии хищников в экосистеме.

Выводы:

Представлена модель типа Лотки–Вольтерры с мутациями — обобщение модели GAN для популяционной генетики.

Обсуждается контроль переобучения для решений этой модели в подходе алгоритмической устойчивости — решения стремятся занять широкие максимумы правдоподобия, узкие пики популяции подавляются (эффект несимметричного отталкивания).

Для популяционной генетики такая модель контроля переобучения объясняет ускорение эволюции при наличии хищников.