Активное обучение

> Reinforcement learning

Прайсинг деривативов с помощью ML

УРМ

Sber

January 15, 2025

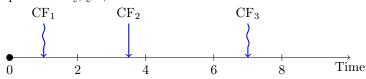
Что такое дериватив?

Подход с линеаризацией

Активное обучение

Reinforcement learning

Это соглашение, по которому стороны обмениваются (возможно случайными) cashflow в некоторые моменты времени в будущем.



Кроме того бывает такое, что одна из сторон может закончить сделку по некоторым условиям в произвольный момент.

Reinforcement learning

В общем случае

 $\mathrm{price} = \mathbb{E}\left(\sum \mathrm{Cashflow}_t \mathcal{DF}(t)\right)$

где \mathcal{DF} — дисконтфактор (возможно стохастический). Таким образом, прайсинг может быть реализован с помощью Монте Карло или биномиальной модели.

- ▶ Это сложно сделать для деривативов общего вида;
- Алгоритмы работают достаточно долго и часто не успевают оценивать целый портфель банка.

Что такое

Подход с линеаризацией

Мы можем разделить состояние рынка и сам контракт дериватива:

- Состояние рынка различные форвардные кривые, волатильности etc (легко векторизуется);
- ▶ Дериватив описание сделки, наиболее общо в json формате.

Можно попробовать сделать 2 модели:

$$\varphi : \text{Markets} \to \mathcal{M} = \mathbb{R}^{n}$$

 $\psi : \text{Derivatives} \to \mathcal{L} = \mathbb{R}^{n}$

Оптимизируем модели

$$\int_{\text{Loans Markets}} \left(\varphi(\mathbf{m}) \cdot \psi^{T}(\mathbf{l}) - \text{price}(\mathbf{m}, \mathbf{l}) \right)^{2} d\mathbf{m} d\mathbf{l} \to 0$$

Remark

Естественно мы считаем, что доступна некоторая (не очень быстрая) функция price, которая может прайсить любой дериватив на любом рынке алгоритмически.

Reinforcement learning

- Необходимо исследовать, как доработать подход так, чтобы за φ и ψ стояли интерпретируемые понятия из финансовой математики репликация сложного дериватива через линейную комбинацию более простых (отсюда и название этого подхода).
- Архитектура универсальна: ее можно сочетать с активным обучением и/или с рекуррентными моделями для эмбеддинга дериватива

В идеале было бы получить следующую архитектуру инференса:

- ► Кроме цены деритватива модель предсказывает и свою уверенность в этой оценке (подобно тому, как устроен uncertainty в Gaussian Process Regressor);
- Таким образом, если модель уверена в своей оценке, то мы будем использовать ее цену;
- Если же неуверенность большая, то мы будем отправлять дериватив считаться алгоритмически и дообучать модель на этой точке.

Активное обучение

Reinforcement learning

В финансах хорошо известны два применения RL, однако они недоисследованы:

- Возможность досрочного погашения приводит к задаче оптимального управления: ее часто решают с помощью техник Монте Карло, но интересно решить более универсальным методом — RL;
- Управление транзакционными издержками также формулируется, как задача оптимального управления.

Remark

Для исследования этих подходов требуется более глубокое понимание финансовой математики.