Игровая модель управления портфелем инвестиций.

Белявский Г. И., Данилова Н. В., Угольницкий Г. А.
(Южный Федеральный университет, г. Ростов-на-Дону)

В докладе задача об оптимальном портфеле рассматривается как игра между двумя игроками, один из которых занимается инвестициями (паевой инвестиционный фонд - ПИФ), другой выделяет на это средства (агент). Основная цель данного подхода – получить условия устойчивого развития изучаемой системы из двух игроков за счет формирования портфеля инвестиций. Для достижения этой цели находится решение игры и изучается задача машинного обучения, с помощью которого реализуется данное решение.

В докладе анализируется литература, связанная с использованием алгоритмов обучения с учителем в реальном времени при формировании портфеля инвестиций, и делается вывод, что данный метод является новым и позволяет получить обоснованные постановки задач обучения для каждого из игроков.

В простом варианте игры имеется один агент, располагающий некоторым резервным капиталом a. Агент делит свой капитал между инвестициями в ПИФ и иными вложениями (например, депозитными вкладами). Получив инвестиции от агента в размере x, ПИФ вкладывает эти средства в ценные бумаги по цене S_0^i, $i = 1,2,\ldots,l$, а именно, $x = \sum_{i=1}^{l} y_i S_0^i$. Получив средства от продажи ценных бумаг по цене $S_1^i = r_i S_o^i$, а именно, $X = \sum_{i=1}^{l} y_i r_i S_0^i$, ПИФ возвращает часть из них αX , и часть $(1-\alpha)X$ оставляет себе, $0 \leq \alpha \leq 1$. Доходность портфеля $Q = \sum_{i=1}^{l} z_i r_i$ является случайной величиной. Доходность агента $K = \alpha Q$. Оптимальная стратегия агента: $x^* = aI_A$, $A = \{\omega: K \geq g\}$. Проблема агента заключается в том, что ему неизвестно значение индикатора. Информация, доступная агенту в момент t – это последовательность $K_1, K_2, \ldots, K_t$, используя которую агент может либо вычислить оценку индикатора, либо вычислить оценку математического ожидания $EI_A = P(A)$, либо оценку среднего EK . Каждую из этих оценок агент вычисляет, применяя обучение в реальном времени. Для каждого из этих регламентов агент выбирает оптимальную стратегию.

Утверждение 1. Для первого информационного регламента оптимальная стратегия агента: $x^* = a\overline{I}_A$. Для второго регламента оптимальная стратегия – смешанная стратегия x^* : $P(x^* = a) = \overline{P}(A), P(x^* = 0) = 1 - \overline{P}(A)$. Для третьего регламента оптимальная стратегия $x^* = \begin{cases} a, \overline{EK} \geq g \\ 0, \overline{EK} < g \end{cases}$.

ПИФ заинтересован в поступлении инвестиций, он стремится добиться этого, выбирая портфель Z и α , поэтому для ПИФа рассматриваются две возможные задачи выбора оптимального решения. Первая задача заключается в максимизации по Z вероятности $P\big((R\,,Z) \geq g/\alpha\big)$ при ограничениях $(I,Z) = 1, (ER,Z) \geq g/\alpha$ и фиксированном α . Первое ограничение непосредственно вытекает из определения доходности портфеля, второе ограничение связано с возможностью использования агентом третьей стратегии. Вторая задача заключается в максимизации по Z порога θ при ограничениях $P\big((Z,R) \geq \theta\big) \geq \beta, (I,Z) = 1, (ER,Z) \geq \theta,\ \theta \geq g$ и фиксированном β. Проблема ПИФа заключается в том, что закон распределения либо известен с точностью до параметров, либо неизвестен совсем. Информация, известная агенту в момент времени t — это последовательность $R_1, R_2, \dots R_t$. В докладе предполагается, что условным законом распределения доходности относительно последовательности является нормальный закон распределения $N\Big(\big(\overline{R}_t, Z\big), \big(\overline{C}_t Z, Z\big)\Big)$, $\overline{R}_t = \frac{1}{t}\sum_{i=1}^{t} R_i$, $\overline{C}_t = \frac{1}{t}\sum_{i=1}^{t} R_i R_i^T - \overline{R}\,\overline{R}^t$, или эмпирический закон распределения, порождаемый последовательностью: $P(R \in A) = \frac{1}{t}\sum_{i=1}^{t} I_{\{R_i \in A\}}$. Для первого варианта, относительно наблюдаемой последовательности, задача максимизации вероятности преобразуется в задачу: $max\,\frac{(Z,\bar{R}_t) - g/\alpha}{\sqrt{(\bar{C}_t Z, Z)}}$ при ограничениях: $(I,Z) = 1, (Z,\bar{R}_t) \geq g/\alpha$. Предполагается, что матрица $\overline{C}_t$ симметричная и положительно определенная. Доказываются следующие утверждения.

Утверждение 2. Если $\bar{R} \neq \mu I$, то задача сводится к одной из двух задач одномерной оптимизации $max\,\frac{\bar{a} + \bar{B}}{\sqrt{\bar{a}^2 + \bar{H}}}$, при ограничении $\bar{a} + \bar{B}\ \geq 0$; или $max\,\frac{-\bar{a} + \tilde{B}}{\sqrt{\bar{a}^2 + \tilde{H}}}$ при ограничении $-\bar{a} + \tilde{B} \geq 0$. Следующее утверждение касается существования решения задачи о максимуме вероятности.

Утверждение 3. Если $\bar{R} \neq \mu I$, то задача максимизации вероятности имеет решение тогда и только тогда, когда либо $\bar{B} > 0$, либо $\tilde{B} > 0$.

Если задача не имеет решения, то критерию следует добавить слагаемое, налагающее штраф за рост нормы вектора Z. Таким образом, первую задачу ПИФа лучше сформулировать следующим образом: $max\left(\frac{(Z,\bar{R}_t) - \theta}{\sqrt{(\bar{C}_t Z, Z)}} - \delta(Z,Z)\right)$ при ограничениях: $(I,Z) = 1, (Z,\bar{R}_t) \geq \theta,\ \delta > 0$. Легко показывается, что данная задача всегда имеет решение.

Если $\overline{R} = \mu I, \mu \geq g/\alpha$, то оптимальное решение $Z^* = \frac{1}{\left(\left(\overline{C}\right)^{-1} I, I\right)}\left(\overline{C}\right)^{-1} I$.

В первом информационном регламенте вторая задача ПИФа будет иметь вид: $max\left[(Z,\bar{R}) - \gamma\sqrt{(\bar{C}Z,Z)}\right]$, при ограничениях $(Z,I) = 1,\ (Z,\bar{R}) - \gamma\sqrt{(\bar{C}Z,Z)} \geq g$. Множитель γ

однозначно определяется из равенства: $\Phi(\gamma) = \beta, \beta \geq 0.5$, Φ - функция Лапласа. Ограничение $(Z, \bar{R}_t) \geq \theta$ выполняется при $\beta \geq 0.5$. Целевая функция данной задачи является вогнутой, так как $\gamma \geq 0$. Рассмотрим задачу без второго ограничения и Z^* - решение этой задачи. Справедливо утверждение.

Утверждение 4. Вторая задача ПИФа имеет решение тогда и только тогда, когда $(Z^*, \bar{R}) - \gamma\sqrt{(\bar{C}Z^*, Z^*)} \geq g$.

Рассмотрим применение ПИФом эмпирического закона распределения, порождаемого последовательностью $R_1, R_2, \dots R_t$. Для эмпирической функции распределения вероятность $P\big((Z, R) \geq \theta\big) = 1 - \frac{Mis}{t}$, $Mis = \sum_{i=1}^{t} I_{\{(Z,R_i)<\theta\}}$, поэтому задача ПИФа заключается в следующем: требуется найти $min\, Mis$ при ограничениях: $(Z, I) = 1, (Z, \bar{R}) \geq g/\alpha$. К сожалению, функции $I_{\{(Z,R_i)<\theta\}}$ не являются выпуклыми, поэтому воспользуемся выпуклой оценкой сверху данных функций: $I_{\{(Z,R_i)<\theta\}} \leq max\{\theta + 1 - (Z, R_i), 0\}$ и рассмотрим выпуклую задачу $min \sum_{i=1}^{t} max\{\theta + 1 - (Z, R_i), 0\}$ при ограничениях: $(Z, I) = 1, (Z, \bar{R}) \geq \theta$. Данную задачу будем рассматривать как первую задачу ПИФа, которую можно решать как задачу обучения с учителем в реальном времени.

Для второй задачи ПИФа воспользуемся предыдущим неравенством, из которого следует, что решение задачи $\max\theta$ при ограничениях: $\frac{1}{t}\sum_{i=1}^{t} max\{\theta + 1 - (Z, R_i), 0\} \leq 1 - \beta, (Z, I) = 1, \big(Z, \overline{R}\big) \geq \theta, \theta \geq g$, является оценкой снизу для решения второй задачи ПИФа.

Как уже отмечалось, задача агента – это задача прогноза, для решения которой естественно использовать обучение с учителем в реальном времени.

Например, прогноз вероятности случайного события, которую агент применяет в своем смешанном решении, может выглядеть следующим образом. Допустим прогноз выражается равенством: $P_{t+1}(A) = \varphi\big((a, \delta_{t-m}^t)\big)$, φ – функция активации со значениями в интервале $[0,1]$, a – параметр, подлежащий определению в результате обучения, $\delta_i = I_{\{K_i \geq g/\alpha\}}$, $\delta_{t-m}^t = \delta_{t-m}, \delta_{t-m+1}, \dots, \delta_t$, $(.,.)$ – скалярное произведение. Качество оценки - $\psi_{t+1}(a)$ определяется после измерения δ_{t+1} следующим образом: $\psi_{t+1}(a) = \delta_{t+1} ln \frac{\varphi\left((a,\delta_{t-m}^t)\right)}{1-\varphi\left((a,\delta_{t-m}^t)\right)} - ln\Big(1 - \varphi\big((a, \delta_{t-m}^t)\big)\Big)$. Для вычисления следующего приближения можно применять различные алгоритмы обучения в реальном времени. Например, уравнение $a_{t+1} = argmin[\|a_t - a\|^2 - h_t\psi_t(a)]$ получается в результате применения алгоритма зеркального спуска.

В завершении доклада приводятся и комментируются результаты вычислений на реальных данных; приводится и комментируется список литературы.