1 Модели экспоненциального сглаживания

§1.1 Простое экспоненциальное сглаживание

Предположим, что мы хотим спрогнозировать некоторый временной ряд $(y_t)_{t=1}^T$. Также предположим, что данный ряд не имеет выраженной сезонности или тренда. Самой простой моделью многошагового прогнозирования можно считать наивную:

$$\hat{y}_{T+h|T} = y_T$$

Данная модель хорошо подходит для бенчмарка, но в большинстве случаев (не всегда!) слишком проста для прогнозирования. Как минимум, она никак не учитывает историю до y_T. Попробуем это исправить. Например, можно добавить усреднение всей истории.

$$\hat{y}_{T+h|T} = \frac{1}{T}\sum_{t=1}^{T} y_t$$

Мы добавили зависимость от истории, однако перестарались. Все наблюдения в таком случае будут иметь одинаковый вес. Логично предположить, что наблюдения, близкие к моменту времени T должны иметь больший вес. Например, если в далёком прошлом, близко к моменту времени 1 временной ряд имел выбросы или структурные сдвиги, не хотелось бы придавать этому большой вес. Сама собой напрашивается геометрическая прогрессия с убывающими весами. Зададим параметр $\alpha \in [0, 1]$ как вес наблюдения y_T и будем уменьшать его на вес q.

$$\hat{y}_{T+h|T} = \frac{1}{T}\sum_{t=1}^{T} \alpha q^{t-1} y_t$$

Найдём веса q. Для простоты предположим, что их сумма должна равняться 1.

$$\alpha + q\alpha + q^2\alpha + \ldots + q^{T-1}\alpha = 1$$

Однако полученное для суммы этой прогрессии уравнение будет зависеть от T и решать его не очень удобно:

$$\frac{\alpha(q^T - 1)}{q - 1} = 1$$

Для упрощения предположим, что T велико и воспользуемся бесконечно убывающей геометрической прогрессией:

$$\frac{\alpha}{q - 1} = 1 \Rightarrow q = 1 - \alpha$$

Таким образом мы получим финальную форму модели:

$$\hat{y}_{T+h|T} = \frac{1}{T}\sum_{t=1}^{T} \alpha(1-\alpha)^{t-1} y_t$$

Веса $\alpha(1-\alpha)^{t-1}$ убывают экспоненциально с ростом t, откуда и получила название модель простого экспоненциального сглаживания. Для дальнейших выводов будет полезно выписать эту модель в двух других представлениях.

1.1.1 Взвешенное среднее

Заметим, что

$$\begin{aligned}\hat{y}_{T+h|T} &= \alpha y_T + \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \dots \\ &= \alpha y_T + (1-\alpha)[\alpha y_{T-1} + \alpha(1-\alpha)^{y}_{T-2} + \dots] \\ &= \alpha y_T + (1-\alpha)\hat{y}_{T|T-1}\end{aligned} \tag{1.1}$$

Получается, что наш прогноз можно представить как взвешенное среднее наблюдаемого значения y_T и его прогноза, полученного на предыдущем шаге $\hat{y}_{T|T-1}$.

1.1.2 Модель коррекции ошибок

Сгруппируем последнее выражение относительно α:

$$\begin{aligned}\hat{y}_{T+h|T} &= \alpha y_T + (1-\alpha)\hat{y}_{T|T-1} \\ &= \alpha(y_T - \hat{y}_{T|T-1}) + \hat{y}_{T|T-1} \\ &= \alpha e_T + \hat{y}_{T|T-1}\end{aligned} \tag{1.2}$$

Так как прогноз на h шагов вперёд является просто экстраполяцией прогноза на один шаг, упростим нашу запись до $h = 1$

$$\begin{aligned}\hat{y}_{T+1|T} &= \alpha y_T + (1-\alpha)\hat{y}_{T|T-1} \\ &= \alpha(y_T - \hat{y}_{T|T-1}) + \hat{y}_{T|T-1} \\ &= \alpha e_T + \hat{y}_{T|T-1}\end{aligned} \tag{1.3}$$

Из этой записи следует, что прогноз можно представить как коррекцию предыдущего прогноза на его ошибку относительно истинного значения с некоторым коэффициентом.

Параметр α можно подобрать, решив задачу оптимизации

$$\sum_{t=1}^{T} (y_t - \hat{y}_{t|t-1})^2 \to \min_{\alpha}$$

Список литературы

[1] https://wikichi.ru/wiki/Doob_decomposition_theorem