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 прогнозирования можно считать наивную:

$$\hat{y}_{T+1|T} = y_T$$

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.

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

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.

$$\hat{y}_{T+1|T} = \frac{1}{T}\sum_{i=0}^{T-1} \alpha q^i y_{T-i}$$

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:

$$\hat{y}_{T+1|T} = \frac{1}{T}\sum_{i=0}^{T-1} \alpha(1-\alpha)^i y_{T-i}$$

Веса $\alpha(1-\alpha)^{t-1}$ убывают экспоненциально с ростом t, откуда и получила название модель простого экспоненциального сглаживания. Многошаговый прогноз такой модели будет плоским и будет просто повторять одношаговый прогноз:

$$\hat{y}_{T+h|T} = \frac{1}{T}\sum_{i=0}^{T-1}\alpha(1-\alpha)^i y_{T-i}$$

Параметр α можно подобрать, численно решив следующую задачу оптимизации:

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

Для дальнейшего анализа будет полезно рассмотрерь несколько дополнительных форм модели экспоненциального сглаживания.

1.1.1 Модель коррекции ошибок

Сгруппируем последнее выражение относительно α:

$$\begin{aligned}\hat{y}_{T+1|T} &= \alpha y_T + (1-\alpha)\hat{y}_{T|T-1} \\ &= \alpha(y_T - \hat{y}_{T|T-1}) + \hat{y}_{T|T-1} \\ &= \alpha e_T + \hat{y}_{T|T-1}\end{aligned} \tag{1.1}$$

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

1.1.2 Взвешенное среднее

Заметим, что

$$\begin{aligned}\hat{y}_{t+1|t} &= \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \ldots \\ &= \alpha y_t + (1-\alpha)[\alpha y_{t-1} + \alpha(1-\alpha)^y_{t-2} + \ldots] \\ &= \alpha y_t + (1-\alpha)\hat{y}_{t|t-1}\end{aligned} \tag{1.2}$$

Получается, что наш прогноз можно представить как взвешенное среднее наблюдаемого значения y_T и его прогноза, полученного на предыдущем шаге $\hat{y}_{T|T-1}$. Однако надо заметить, что для корректности такой формы нужно ввести один дополнительный параметр l_0, инициализирующий последовательность. Далее станет ясно, почему в качестве имени мы взяли именно l_0.

$$\hat{y}_{2|1} = \alpha y_1 + (1-\alpha)l_0$$

Тогда прогнозное уравнение также изменится.

$$\hat{y}_{T+1|T} = \frac{1}{T}\sum_{i=0}^{T-1}\alpha(1-\alpha)^i y_{T-i} + (1-\alpha)^T l_0$$

Вес последнего слагаемого будет быстро убывать при больших T, и модель будет эквивалентна стандартной постановке. Для полной эквивалентности можно просто положить $l_0 = 0$.

Параметр l_0 можно найти из той же задачи оптимизации:

$$\sum_{t=1}^{T}(y_t - \hat{y}_{t|t-1})^2 \to \min_{\alpha, l_0}$$

1.1.3 Компонентный вид

$$\begin{aligned} &\text{Уравнение прогноза} && \hat{y}_{t+1|t} = l_t \\ &\text{Уравнение сглаживания} && l_t = \alpha y_t + (1-\alpha) l_{t-1} \end{aligned} \tag{1.3}$$

Эта формулировка эквивалентна предыдущим. Она удобна технически для того, чтобы впоследствии добавлять уравнения и новые компоненты в уравнение прогноза. Здесь мы также заострим внимание на том, что l_t в такой постановке можно интерпретировать как сглаженный уровень ряда.

§1.2 Сглаживание с трендом

Предыдущая модель подходит только для данных без ярко выраженных трендов. Для добавления большей динамики введём ещё один показатель. b_t будет означать локальную скорость роста за один период модели. Грубо говоря, этот параметр будет отвечать за приращения компоненты l_t

Список литературы