

Форель (алгоритм прогнозирования)

”Форель” — алгоритм прогнозирования [[временной ряд|временного ряда]] с выраженной периодической составляющей. Для прогноза использует [[скрытая переменная|скрытые регрессионные переменные]] и [[корректирующая операция|корректирующие операции]] над множеством выбираемых моделей.

tip|Авторы рекомендуют посетить [[Обсуждение:Оптимальная хирургия мозга|обсуждение]] до прочтения статьи.

Постановка задачи

Дан временной ряд $\mathbf{x} = [x_1, \dots, x_T]^T$, $x \in \mathbb{R}^1$. * Предполагается, что отсчеты времени сделаны через равные промежутки. Следовательно, отсчеты времени t можно без потери общности заменить на индексы — элементы натурального ряда. * Предполагается, что ряд имеет периодическую составляющую, и длина периода k известна. * Предполагается, что ряд, возможно, имеет пропущенные значения. * Предполагается, что длина ряда кратна периоду. Это условие удовлетворяется присоединением к началу ряда необходимого числа пропущенных значений.

Требуется построить [[Алгоритм обучения|алгоритм]], выполняющий прогноз с [[горизонт прогноза|горизонтом прогноза]] равным длине одного периода.

Описание алгоритма

Составляется $(m \times k)$ -таблица значений временного ряда:

$$X = \begin{pmatrix} x_T & x_{T-1} & \dots & x_{T-k+1} \\ x_{(m-1)k} & x_{(m-1)k-1} & \dots & x_{(m-2)k+1} \\ \dots & \dots & \dots & \dots \\ x_{nk} & x_{nk-1} & \dots & x_{n(k-1)+1} \\ \dots & \dots & \dots & \dots \\ x_k & x_{k-1} & \dots & x_1 \end{pmatrix}$$

в которой длина ряда $T = mk$. Каждый столбец ряда содержит элементы с индексами, кратными периоду.

Пример. Если выполняется почасовой прогноз с горизонтом одна неделя (по условию задачи период также равен одной неделе), то $k = 24 \times 7 = 168$. При этом в первой строке окажутся значения временного ряда за последнюю неделю, начиная с самого позднего отсчета (пусть это будет пятница, 12 часов). В первом столбце окажутся все значения, которые соответствуют пятницам, 12 часов; верхний элемент столбца будет самым поздним по времени, а нижний будет соответствовать первой пятнице 12 часов ряда. Если считать временной ряд гармоническим, то элементы одного столбца будут иметь одну и ту же фазу.

Обозначим столбцы матрицы $\mathbf{x}_k, \dots, \mathbf{x}_1$. Для каждого столбца i матрицы X построим набор моделей-предикторов. Для этого зафиксируем столбец \mathbf{x}_i , считая, что прогнозируем значение ряда в момент времени $i + k$.

Для каждого из прочих столбцов $\mathbf{x}_j, j = 1, \dots, k$, и $j \neq i$ решим задачу [[метод наименьших квадратов|линейной регрессии]] $\|\mathbf{x}_i - G_j \mathbf{w}\|^2 \rightarrow \min$, где матрица

$$G_j = \begin{pmatrix} g_1(x_{mj}) & g_2(x_{mj}) & \dots & g_r(x_{mj}) \\ g_1(x_{(m-1)j}) & g_2(x_{(m-1)j}) & \dots & g_r(x_{(m-1)j}) \\ \dots & \dots & \dots & \dots \\ g_1(x_j) & g_2(x_j) & \dots & g_r(x_j) \end{pmatrix}.$$

Функции g_1, \dots, g_r заданы или определены исходя из дополнительных условий, см. например методы [[символьная регрессия|символьной регрессии]].

Пример. Модель-предиктор имеет вид

$$f(x) = w_1 + w_2\sqrt{x} + w_3x + w_4x\sqrt{x}.$$

Соответственно, $g_1 = 1$, $g_2 = \sqrt{x}$, $g_3 = x$, $g_4 = x\sqrt{x}$ и $\mathbf{w} = [w_1, w_2, w_3, w_4]^T$.

Выбирается заданное число p векторов $G_j\mathbf{w}$, доставляющих наибольшее значение функционалу $\rho(\mathbf{x}_i, G_j\mathbf{w})$. Обозначим P — множество выбранных индексов $\{j\}$.

Пример. Функционал $\rho(\mathbf{x}_1, \mathbf{x}_2)$ задан как коэффициент корреляции двух векторов; задан как норма разности двух векторов.

Строится корректор над множеством моделей-предикторов — линейная регрессия $\|\mathbf{x}_i - H_P\mathbf{b}\|^2 \rightarrow \min$ с ограничением на неотрицательность вектов \mathbf{b} . Матрица H_P — присоединенные векторы $G_j\mathbf{w}$, индексы $j \in P$.

Прогнозируемое значение ряда \mathbf{x} в момент времени $mk+i$ равно значению первого элемента вектора $H_P\mathbf{b}$.

Смотри также

- * [[Метод наименьших квадратов]]
- * [[Символьная регрессия]]
- * [[Регрессионный анализ]]
 - [[Категория:Регрессионный анализ]]
 - [[Категория:Прогнозирование]]