Добавляем предикторы

Как обойтись без моделей?

Обойтись без моделей: план

- Как переделать временные ряды в перекрестные данные?

Обойтись без моделей: план

- Как переделать временные ряды в перекрестные данные?
- Добавить лаги переменной y_t.

Обойтись без моделей: план

- Как переделать временные ряды в перекрестные данные?
- Добавить лаги переменной y_t.
- Использовать агрегирующие функции и скользящее или растущее окно.

Как обойтись без моделей?

Старые друзья

Есть алгоритмы, которые по обучающей выборке зависимой переменной y, обучающей матрице предикторов X, и новым предикторам X_F строят прогноз $\hat{y}_F$.

Как обойтись без моделей?

Старые друзья

Есть алгоритмы, которые по обучающей выборке зависимой переменной y, обучающей матрице предикторов X, и новым предикторам X_F строят прогноз $\hat{y}_F$.

Случайный лес, градиентный бустинг...

Как обойтись без моделей?

Старые друзья

Есть алгоритмы, которые по обучающей выборке зависимой переменной y, обучающей матрице предикторов X, и новым предикторам X_F строят прогноз $\hat{y}_F$.

Случайный лес, градиентный бустинг... и даже обычная регрессия!

Как обойтись без моделей?

Старые друзья

Есть алгоритмы, которые по обучающей выборке зависимой переменной y, обучающей матрице предикторов X, и новым предикторам X_F строят прогноз $\hat{y}_F$.

Случайный лес, градиентный бустинг… и даже обычная регрессия!

Можно усреднять прогнозы ARIMA/ETS и прогнозы других алгоритмов.

Как создать предикторы?

Из одного столбца y можно создать целую матрицу X предикторов!

- Использовать лаги y_{t-k}.

Как создать предикторы?

Из одного столбца y можно создать целую матрицу X предикторов!

- Использовать лаги y_{t-k}.
- Использовать функции от лагов в качестве предикторов.

Используем лаги y

Для примера возьмём два лага, Ly_t и L^2y_t.

Используем лаги y

Для примера возьмём два лага, Ly_t и L^2y_t.

Обучающая выборка:

$$\begin{pmatrix} y_3 \\ y_4 \\ y_5 \\ \vdots \\ y_T \end{pmatrix} \qquad \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \\ y_3 & y_4 \\ \vdots & \vdots \\ y_{T-2} & y_{T-1} \end{pmatrix}$$

Используем лаги y

Для примера возьмём два лага, Ly_t и L^2y_t.

Обучающая выборка:

$$\begin{pmatrix} y_3 \\ y_4 \\ y_5 \\ \vdots \\ y_T \end{pmatrix} \quad \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \\ y_3 & y_4 \\ \vdots & \vdots \\ y_{T-2} & y_{T-1} \end{pmatrix}$$

Выборка для прогнозирования:

$$\begin{pmatrix} ? \end{pmatrix} \quad \begin{pmatrix} y_{T-1} & y_T \end{pmatrix}$$

Сколько лагов добавить?

- Каждый добавленный лаг сокращает обучающую выборку!

Сколько лагов добавить?

- Каждый добавленный лаг сокращает обучающую выборку!
- Разумно добавить ближайшие лаги Ly_t, L^2y_t.

Сколько лагов добавить?

- Каждый добавленный лаг сокращает обучающую выборку!
- Разумно добавить ближайшие лаги Ly_t, $L^2 y_t$.
- Для сезонных данных разумно добавить сезонный лаг $L^{12} y_t$.

Сколько лагов добавить?

- Каждый добавленный лаг сокращает обучающую выборку!
- Разумно добавить ближайшие лаги Ly_t, L^2y_t.
- Для сезонных данных разумно добавить сезонный лаг $L^{12}y_t$.
- Есть алгоритмы чувствительные к лишним предикторам: например, регрессия.

Сколько лагов добавить?

- Каждый добавленный лаг сокращает обучающую выборку!
- Разумно добавить ближайшие лаги Ly_t, L^2y_t.
- Для сезонных данных разумно добавить сезонный лаг $L^{12}y_t$.
- Есть алгоритмы чувствительные к лишним предикторам: например, регрессия.
- Есть алгоритмы нечувствительные к лишним предикторам: например, случайный лес.

Функции от лагов

При прогнозировании y_t честно использовать любую функцию от предыдущих $y_{t-1}, y_{t-2}, \ldots$

Функции от лагов

При прогнозировании y_t честно использовать любую функцию от предыдущих $y_{t-1}, y_{t-2}, \ldots$

Например:

- $\Delta y_{t-1} = y_{t-1} - y_{t-2}$;

Функции от лагов

При прогнозировании y_t честно использовать любую функцию от предыдущих $y_{t-1}, y_{t-2}, \dots$

Например:

- $\Delta y_{t-1} = y_{t-1} - y_{t-2}$;
- $\max\{y_{t-1}, y_{t-2}, y_{t-3}\}$;

Функции от лагов

При прогнозировании y_t честно использовать любую функцию от предыдущих $y_{t-1}, y_{t-2}, \dots$

Например:

- $\Delta y_{t-1} = y_{t-1} - y_{t-2}$;
- $\max\{y_{t-1}, y_{t-2}, y_{t-3}\}$;
- $\min\{y_{t-1}, y_{t-2}, \dots, y_1\}$.

Типичный предиктор

- Агрегирующая функция:
  Минимум, максимум, среднее, медиана, размах, выборочная дисперсия, выборочное стандартное отклонение, ...

Типичный предиктор

- Агрегирующая функция:
  Минимум, максимум, среднее, медиана, размах, выборочная дисперсия, выборочное стандартное отклонение, $\ldots$
- Аргумент агрегирующей функции:
  Скользящее окно: агрегирующая функция применяется, скажем, к трём предыдущим значениям y_{t-1}, y_{t-2}, y_{t-3}.
  Растущее окно: агрегирующая функция применяется ко всем предыдущим значениям y_{t-1}, y_{t-2}, $\ldots$, y_1.

Используем функции лагов y

Для примера возьмём максимум скользящим окном и минимум растущим окном.

Используем функции лагов y

Для примера возьмём максимум скользящим окном и минимум растущим окном.

Обучающая выборка:

$$
\begin{pmatrix} y_3 \\ y_4 \\ y_5 \\ \vdots \\ y_T \end{pmatrix}
\qquad
\begin{pmatrix}
\max\{y_1, y_2\} & \min\{y_1, y_2\} \\
\max\{y_2, y_3\} & \min\{y_1, y_2, y_3\} \\
\max\{y_3, y_4\} & \min\{y_1, \ldots, y_4\} \\
\vdots & \vdots \\
\max\{y_{T-2}, y_{T-1}\} & \min\{y_1, \ldots, y_{T-1}\}
\end{pmatrix}
$$

Используем функции лагов y

Для примера возьмём максимум скользящим окном и минимум растущим окном.

Обучающая выборка:

$$\begin{pmatrix} y_3 \\ y_4 \\ y_5 \\ \vdots \\ y_T \end{pmatrix} \quad \begin{pmatrix} \max\{y_1, y_2\} & \min\{y_1, y_2\} \\ \max\{y_2, y_3\} & \min\{y_1, y_2, y_3\} \\ \max\{y_3, y_4\} & \min\{y_1, \dots, y_4\} \\ \vdots & \vdots \\ \max\{y_{T-2}, y_{T-1}\} & \min\{y_1, \dots, y_{T-1}\} \end{pmatrix}$$

Выборка для прогнозирования:

$$\begin{pmatrix} ? \end{pmatrix} \quad \begin{pmatrix} \max\{y_{T-1}, y_T\} & \min\{y_1, \dots, y_T\} \end{pmatrix}$$

Обойтись без моделей: итоги

- Помните о случайном лесе, градиентном бустинге и даже об обычной регрессии.

Обойтись без моделей: итоги

- Помните о случайном лесе, градиентном бустинге и даже об обычной регрессии.
- Добавьте лаги зависимой переменной.

Обойтись без моделей: итоги

- Помните о случайном лесе, градиентном бустинге и даже об обычной регрессии.
- Добавьте лаги зависимой переменной.
- Добавьте агрегирующие функции скользящим и растущим окном.

У нас есть ещё время!

У нас ещё есть время: план

- Предикторы тренда.

У нас ещё есть время: план

- Предикторы тренда.
- Сезонные и праздничные дамми.

У нас ещё есть время: план

- Предикторы тренда.
- Сезонные и праздничные дамми.
- Косинусы и синусы.

Используем время!

Для примера возьмём t и $\sqrt{t}$.

Используем время!

Для примера возьмём t и $\sqrt{t}$.

Обучающая выборка:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_T \end{pmatrix} \qquad \begin{pmatrix} 1 & \sqrt{1} \\ 2 & \sqrt{2} \\ 3 & \sqrt{3} \\ \vdots & \vdots \\ T & \sqrt{T} \end{pmatrix}$$

Используем время!

Для примера возьмём t и $\sqrt{t}$.

Обучающая выборка:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_T \end{pmatrix} \qquad \begin{pmatrix} 1 & \sqrt{1} \\ 2 & \sqrt{2} \\ 3 & \sqrt{3} \\ \vdots & \vdots \\ T & \sqrt{T} \end{pmatrix}$$

Выборка для прогнозирования:

$$\begin{pmatrix} ? \end{pmatrix} \qquad \begin{pmatrix} T+1 & \sqrt{T+1} \end{pmatrix}$$

Включать ли монотонные преобразования времени?

- Всегда можно попробовать включить!

Включать ли монотонные преобразования времени?

- Всегда можно попробовать включить!
- Алгоритмам основанным на построении деревьев (случайные лес, градиентный бустинг) дополнительные монотонные преобразования времени бесполезны.

Включать ли монотонные преобразования времени?

- Всегда можно попробовать включить!
- Алгоритмам основанным на построении деревьев (случайные лес, градиентный бустинг) дополнительные монотонные преобразования времени бесполезны.
- Помните о возможном преобразовании исходной переменной (логарифм, преобразование Бокса-Кокса).

Сезонные и праздничные дамми

Если сезонов немного, то разумно включить дамми на каждый сезон.

Сезонные и праздничные дамми

Если сезонов немного, то разумно включить дамми на каждый сезон.

Обучающая выборка для квартальных данных:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ \vdots \\ y_T \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Ловушка дамми-переменных

В регрессии помните о ловушке дамми-переменных!

Ловушка дамми-переменных

В регрессии помните о ловушке дамми-переменных!

- Либо дамми на каждый сезон и модель без константы.

Ловушка дамми-переменных

В регрессии помните о ловушке дамми-переменных!

- Либо дамми на каждый сезон и модель без константы.
- Либо дамми на все сезоны кроме одного и модель с константой.

Ловушка дамми-переменных

В регрессии помните о ловушке дамми-переменных!

- Либо дамми на каждый сезон и модель без константы.
- Либо дамми на все сезоны кроме одного и модель с константой.

Ловушка дамми-переменных

В регрессии помните о ловушке дамми-переменных!

- Либо дамми на каждый сезон и модель без константы.
- Либо дамми на все сезоны кроме одного и модель с константой.

Алгоритмы основанные на построении деревьев (случайные лес, градиентный бустинг) устойчивы к ловушке дамми.

Зачем нужны синусы и косинусы?

Стратегия добавления всех дамми переменных плохо работает, если их нужно много.

Зачем нужны синусы и косинусы?

Стратегия добавления всех дамми переменных плохо работает, если их нужно много.

Вряд ли стоит добавлять 365 дамми-переменных для дневных данных.

Зачем нужны синусы и косинусы?

Стратегия добавления всех дамми переменных плохо работает, если их нужно много.

Вряд ли стоит добавлять 365 дамми-переменных для дневных данных.

Обойтись малым числом предикторов помогут синус и косинус!

Зачем нужны синусы и косинусы?

Стратегия добавления всех дамми переменных плохо работает, если их нужно много.

Вряд ли стоит добавлять 365 дамми-переменных для дневных данных.

Обойтись малым числом предикторов помогут синус и косинус!

Два факта:

- Период у $\sin t$ и $\cos t$ равен 2π;

Зачем нужны синусы и косинусы?

Стратегия добавления всех дамми переменных плохо работает, если их нужно много.

Вряд ли стоит добавлять 365 дамми-переменных для дневных данных.

Обойтись малым числом предикторов помогут синус и косинус!

Два факта:

- Период у $\sin t$ и $\cos t$ равен 2π;
- При умножении аргумента на a период сокращается в a раз.

Разложение Фурье

Теорема

Любая непрерывная и дифференциируемая функция f с периодом 2π может быть представлена в виде

$$f(t) = c + \sum_{k=1}^{\infty} a_k \cos(kt) + b_k \sin(kt).$$

Разложение Фурье

Теорема

Любая непрерывная и дифференциируемая функция f с периодом 2π может быть представлена в виде

$$f(t) = c + \sum_{k=1}^{\infty} a_k \cos(kt) + b_k \sin(kt).$$

Практический рецепт для дневных данных:

- Добавьте предикторы $\cos\left(\frac{2\pi}{365} \cdot t\right)$ и $\sin\left(\frac{2\pi}{365} \cdot t\right)$;

Разложение Фурье

Теорема

Любая непрерывная и дифференциируемая функция f с периодом 2π может быть представлена в виде

$$f(t) = c + \sum_{k=1}^{\infty} a_k \cos(kt) + b_k \sin(kt).$$

Практический рецепт для дневных данных:

- Добавьте предикторы $\cos\left(\frac{2\pi}{365} \cdot t\right)$ и $\sin\left(\frac{2\pi}{365} \cdot t\right)$;
- Добавьте предикторы $\cos\left(\frac{2\pi}{365} \cdot 2t\right)$ и $\sin\left(\frac{2\pi}{365} \cdot 2t\right)$;

Разложение Фурье

Теорема

Любая непрерывная и дифференциируемая функция f с периодом 2π может быть представлена в виде

$$f(t) = c + \sum_{k=1}^{\infty} a_k \cos(kt) + b_k \sin(kt).$$

Практический рецепт для дневных данных:

- Добавьте предикторы $\cos\left(\frac{2\pi}{365} \cdot t\right)$ и $\sin\left(\frac{2\pi}{365} \cdot t\right)$;
- Добавьте предикторы $\cos\left(\frac{2\pi}{365} \cdot 2t\right)$ и $\sin\left(\frac{2\pi}{365} \cdot 2t\right)$;
- Добавьте предикторы $\cos\left(\frac{2\pi}{365} \cdot 3t\right)$ и $\sin\left(\frac{2\pi}{365} \cdot 3t\right)$;

Разложение Фурье

Теорема

Любая непрерывная и дифференциируемая функция f с периодом 2π может быть представлена в виде

$$f(t) = c + \sum_{k=1}^{\infty} a_k \cos(kt) + b_k \sin(kt).$$

Практический рецепт для дневных данных:

- Добавьте предикторы $\cos\left(\frac{2\pi}{365} \cdot t\right)$ и $\sin\left(\frac{2\pi}{365} \cdot t\right)$;
- Добавьте предикторы $\cos\left(\frac{2\pi}{365} \cdot 2t\right)$ и $\sin\left(\frac{2\pi}{365} \cdot 2t\right)$;
- Добавьте предикторы $\cos\left(\frac{2\pi}{365} \cdot 3t\right)$ и $\sin\left(\frac{2\pi}{365} \cdot 3t\right)$;
- $\ldots$

У нас есть ещё время: итоги

- Используйте время в качестве предиктора.

У нас есть ещё время: итоги

- Используйте время в качестве предиктора.
- Сезонность в предикторах можно отразить с помощью дамми-переменных или с помощью косинуса и синуса.