

Выбор суперпозиции моделей при прогнозировании грузовых железнодорожных перевозок

Двинских Дарина Михайловна

Московский физико-технический институт
Факультет управления и прикладной математики
Кафедра «Интеллектуальные системы»

Научный руководитель д.ф.-м.н., В. В. Стрижов

2016 г.

Задача

Построить суперпозицию моделей для повышения эффективности транспортировки грузов

Проблема

- Отсутствие единой прогностической модели
- Большая волатильность временных рядов
- Наличие нулевых объемов перевозок

Требования к моделям

- Обоснованность (соответствие цели прогноза)
- Валидность (способность учитывать изменение ряда)
- Точность (способность давать оценку параметров, минимально отличающуюся от их фактического значения.)

- *К.В. Рудаков, М.П. Кузнецов, А.П. Мотренко, М.М. Стенина, Д.О.Каширин, В.В. Стрижов*, Выбор оптимальной модели прогнозирования грузовых железнодорожных перевозок, 2015
- *Б.Б. Егамбердиев, М.Н. Витченко*, Прогнозирования объемов международных грузовых перевозок на железнодорожном транспорте с использованием приемов эконометрики, 2012
- *Д.В. Железнов, А.П. Батурин* Методология усиления провозной способности железных дорог России в условиях реформы отрасли, 2013

дата погрузки	код станции отправления	код станции назначения	код груза	суммарный вес груза
2007-01-01	14605	831504	1	56
2007-01-01	135602	165504	18	63
2007-01-01	830304	814208	13	246

- **Дата погрузки:** годовая история перевозок
- **Коды станций отправления и назначения:** 78 регионов, около 4000 станций
- **Код груза:** 43 наименования - нефть, руда, торф, автомобили, хлопок, сахар, зерно, ...

Задача построения оптимальной суперпозиции

Рассмотрим временной ряд $\mathbf{x} = \{x_i\}_{i=1}^T$ из системы временных рядов.

Пусть f и g - две базовые функции прогнозирования, причем функция f прогнозирует ряд, а g - его остатки

$$f : x_t \rightarrow \hat{x}_{t+1}^f, \quad g : x_{t+1}^f \rightarrow \hat{x}_{t+1}^{f,g}.$$

Предположим, что значения временного ряда, можно приблизить суперпозицией с допустимой точностью

$$x_{t+1} = f \circ g(x_t, x_{t-1}, \dots, x_1) + \epsilon_{t+1}.$$

Требуется построить прогноз

$$\hat{x}_{T+1} = f \circ g(\hat{\mathbf{w}}, x_T, x_{T-1}, \dots, x_1)$$

и ретроспективный прогноз с горизонтом прогнозирования h

$$\hat{x}_{t+h} = f \circ g(\hat{\mathbf{w}}, x_t, x_{t-1}, \dots, x_1).$$

Функция ошибки $S(\mathbf{w}|F, D)$ — функция, значение которой требуется минимизировать для получения оценок параметров $\hat{\mathbf{w}}$ модели

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w} \in \mathcal{W}} \sum_{t \in \mathcal{T}} S(\mathbf{w}, x_t),$$

$$\epsilon = x_t - \hat{x}_t,$$

где $\hat{x}_t = f \circ g(\hat{\mathbf{w}}, x_t, x_{t-1}, \dots, x_1)$.

$$MAE = \frac{1}{n} \sum_{i=1}^n |\epsilon_i|$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\epsilon_i}{x_i} \right|$$

$$PMAD = \sum_{i=1}^n |\epsilon_i| (\sum_{i=1}^n |x_i|)^{-1}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n \epsilon_i^2$$

$$RMSE = \frac{1}{\sqrt{n}} \sqrt{\sum_{i=1}^n \epsilon_i^2}$$

$$SS = 1 - \frac{MSE_{forecast}}{MSE_{history}}$$

Рассмотрим функции из допустимого семейства моделей \mathcal{F} .

Наивный прогноз

$$\hat{x}_{T+1} = x_T.$$

Прогноз средним значением за последние k отсчетов

$$\hat{x}_{T+1} = \sum_{t=T-k+1}^T x_t.$$

Скользящее среднее

$$z_t = \frac{x_t + x_{t+1} + \dots + x_{t+n-1}}{n} = \sum_{i=0}^{n-1} x_{t+i},$$

где n — ширина окна,

$$\hat{x}_{T+h} = z_T.$$

Экспоненциальное сглаживание

$$z_t = \alpha x_t + (1 - \alpha)z_{t-1},$$

где x — исходный ряд, z — сглаженный ряд, α — параметр сглаживания ряда, $\alpha \in (0, 1)$,

$$\hat{x}_{T+h} = z_T.$$

Метод Кростена

$\mathbf{d} = \{d_t\}_{t=1}^T$ — ненулевой спрос исходного временного ряда x ,
 $\mathbf{q} = \{q_t\}_{t=1}^T$ — интервалы между ненулевым спросом ряда x .
 Экспоненциальное сглаживания обоих рядов

$$z_t = \alpha d_t + (1 - \alpha)z_{t-1},$$

$$p_t = \alpha q_t + (1 - \alpha)p_{t-1},$$

$$\hat{x}_{T+h} = \frac{z_T}{p_T}.$$

Модель ARIMA(p, d, q)

$$\nabla^d x_t = c + \sum_{i=1}^p a_i \nabla^d x_{t-i} + \sum_{j=1}^q b_j \epsilon_{t-j} + \epsilon_t,$$

где ϵ_t — стационарный временной ряд,

c, a_i, b_j — параметры модели,

∇^d — оператор разности временного ряда порядка d .

Модель VAR

$$\mathbf{Y} = \mathbf{XW},$$

где матрица \mathbf{X} — матрица объект-признак, \mathbf{W} — матрица весов, \mathbf{Y} — матрица ответов

$$\mathbf{W} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y},$$

тогда вектор прогнозов

$$\hat{\mathbf{x}} = \mathbf{xW}.$$

Свойства временных рядов

Сопоставление различных прогностических моделей свойствам временных рядов

Алгоритмы	Свойства временных рядов				
	сезонность	тренд	нестац.	нулевые значения	внешние факторы
Среднее	-	-	-	+	-
Скользящее среднее	-	-	-	+	-
Сглаживание	-	-	-	+	-
Метод Кростона	-	-	-	+	-
ARIMA	+	+	+	-	-
VAR	+	+	-	-	+

- 1 С помощью базовой функции f вычисляется n прогнозов конца истории $\hat{x}_t^f, \dots, \hat{x}_{t-n+1}^f$ на одну точку.
- 2 Вычисляется n остатков $\hat{\varepsilon}_t, \dots, \hat{\varepsilon}_{t-n+1}$ в виде разницы

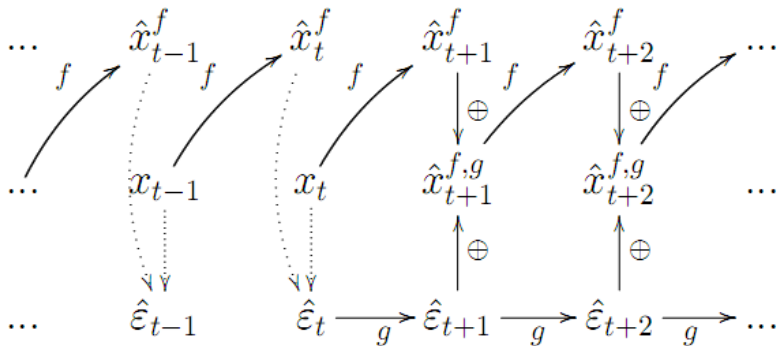
$$\hat{\varepsilon}_{t-k} = x_{t-k} - \hat{x}_{t-k}^f.$$

- 3 С помощью функции g прогнозируются остатки $\hat{\varepsilon}_{t+i}$ на горизонт прогнозирования h .
- 4 Выполняется итеративный подсчет конечных прогнозов

$$\hat{x}_{t+i}^{f,g} = \hat{x}_{t+i}^f + \hat{\varepsilon}_{t+i}$$

с последовательным подсчетом прогноза базовой функцией f на одну точку \hat{x}_{t+i}^f .

Диаграмма прогнозирования исходного ряда и его остатков



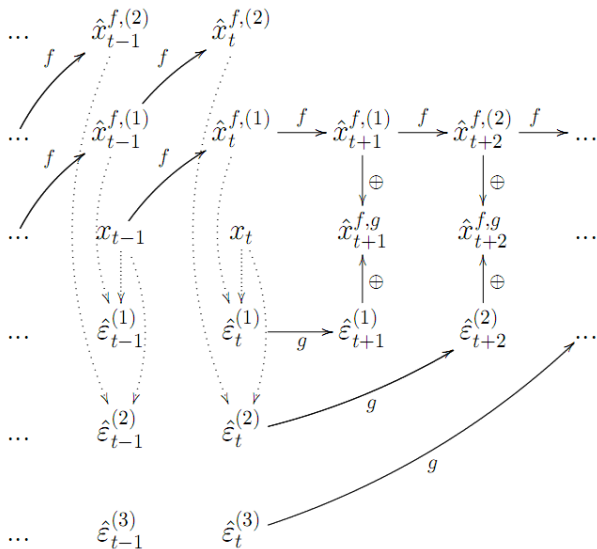
- 1 С помощью базовой функции f вычисляется ретроспективный прогноз $\hat{x}_{t+1}^{f,(1)}, \dots, \hat{x}_{t+i}^{f,(i)}$ с горизонтом прогнозирования i , каждый — на глубине i .
- 2 С помощью базовой функции f вычисляется h наборов прогнозов конца истории, $\hat{x}_t^{f,(i)}, \dots, \hat{x}_{t-n+1}^{f,(i)}$, каждый набор — на глубине i , $i = 1, \dots, h$.
- 3 Вычисляется h наборов остатков $\hat{\varepsilon}_t^{(i)}, \dots, \hat{\varepsilon}_{t-n+1}^{(i)}$

$$\hat{\varepsilon}_{t-k}^{(i)} = x_{t-k} - \hat{x}_{t-k}^{f,(i)}, \quad i = 1, \dots, h.$$

- 4 С помощью функции g прогнозируются остатки $\hat{\varepsilon}_{t+i}^{(i)}$, каждый прогноз выполняется на одну точку и использует вычисленную последовательность $\hat{\varepsilon}_t^{(i)}, \dots, \hat{\varepsilon}_{t-n+1}^{(i)}$.
- 5 Выполняется подсчет конечных прогнозов

$$\hat{x}_{t+i}^{f,g} = \hat{x}_{t+i}^{f,(i)} + \hat{\varepsilon}_{t+i}^{(i)}.$$

Диаграмма прогнозирования исходного ряда и его остатков



Цель

Проверить целесообразность использования суперпозиции для прогнозирования временных рядов

Средства - используемые модели

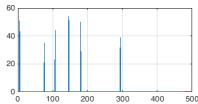
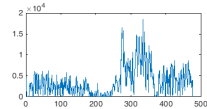
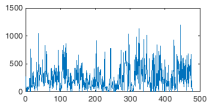
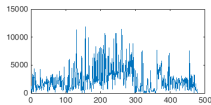
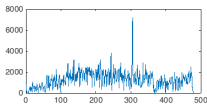
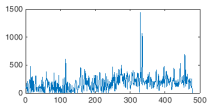
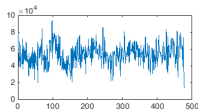
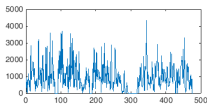
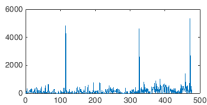
Среднее значение, MA, экспоненциальное сглаживание, метод Кростона, ARIMA, VAR

Данные

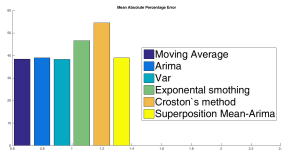
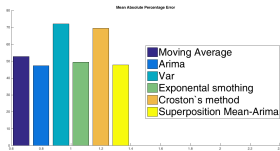
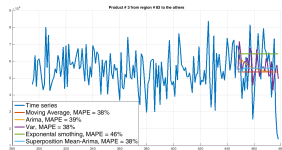
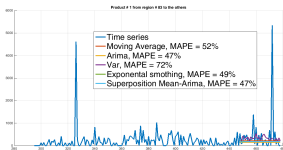
Временные ряды

Вычислительный эксперимент. Получение временных рядов

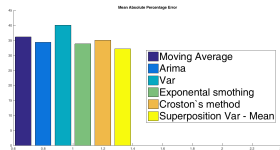
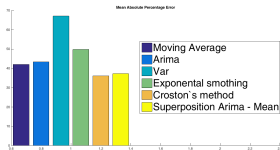
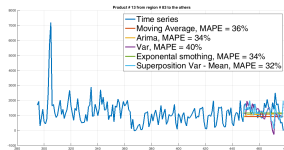
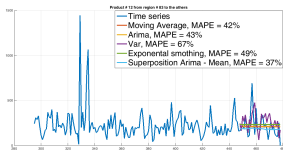
Проведена дополнительная агрегация временных рядов.
Товары, отправленные с 83-го региона: каменный уголь, кокс, нефть, метизы, лом черных металлов, строительные грузы, зерно и бумага.



Вычислительный эксперимент. Сравнение ошибок



Вычислительный эксперимент. Сравнение ошибок



- Выбраны базовые модели для прогнозирования с учетом специфики временных рядов: среднее значение, MA, ARIMA, VAR, экспоненциальное сглаживание, метод Кростона
- Построены суперпозиции моделей: всевозможные комбинации базовых моделей
- Проведен вычислительный эксперимент, сравнивающий качество базовых моделей и построенных суперпозиций
- Продемонстрирована целесообразность использования суперпозиции для прогнозирования временных рядов