

Вариации ETS

Модель ETS: дампированный тренд

Модель ETS: дампированный тренд

- Идея дампированного тренда.
- Новый коэффициент в модели.
- Формулы для прогнозов.

Проблема тренда в ETS(AAN)

В ETS(AAN) модели **скорость роста** тренда ℓ_t определена формулой

$$b_t = b_{t-1} + \beta u_t, \text{ стартовое } b_0.$$

Следовательно,

$$\mathbb{E}(b_t) = \mathbb{E}(b_{t-1}), \quad \mathbb{E}(b_{T+h} \mid b_T) = b_T.$$

Долгосрочный прогноз положительного показателя при $b_T < 0$ станет отрицательным.

Противоречие

Краткосрочные ожидания изменения показателя.

Хотим тренд в модели.

Долгосрочная невозможность отрицательных значений.

Не хотим тренд в модели.

Решение: **дампированный** или **затухающий** тренд.

Лишние параметры — дорого!

Хотим более богатую динамику тренда — нужны **дополнительные** параметры.

Дополнительные параметры — риск **переподгонки** модели, более **широкие доверительные интервалы** для оставшихся параметров.

Обойдёмся всего **одним** новым параметром!

Дампированный тренд

Вводим параметр затухания тренда $\phi \in (0; 1)$ в уравнение наклона:

$$b_t = \phi b_{t-1} + \beta u_t, \text{ стартовое } b_0.$$

И в остальные уравнения:

$$\begin{cases} y_t = \ell_{t-1} + \phi b_{t-1} + u_t; \\ \ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha u_t, \text{ стартовое } \ell_0; \end{cases}$$

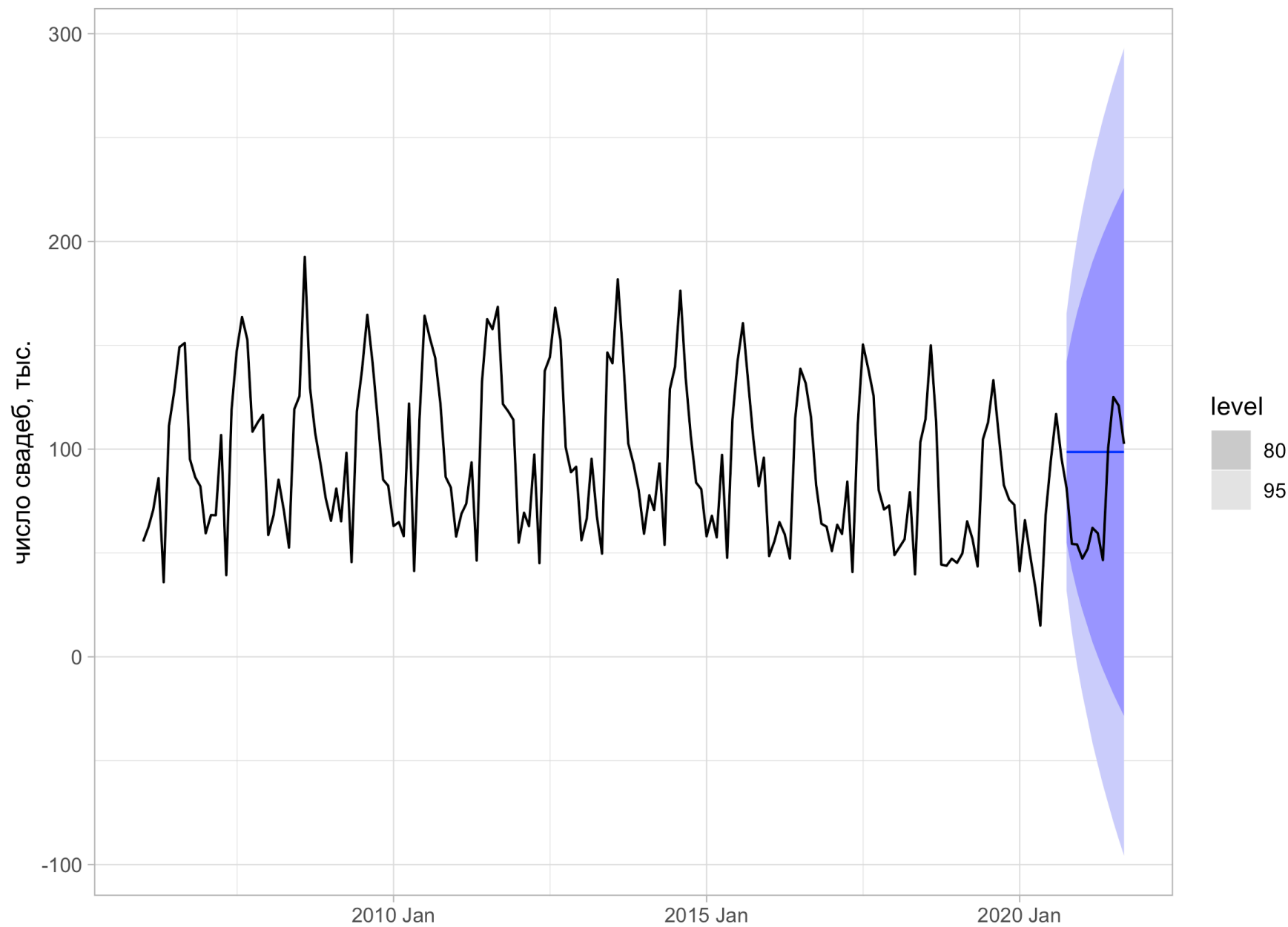
Общий вид ETS(AAdN)

$$\left\{ \begin{array}{l} y_t = \ell_{t-1} + \phi b_{t-1} + u_t; \\ \ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha u_t, \text{ стартовое } \ell_0; \\ b_t = \phi b_{t-1} + \beta u_t, \text{ стартовое } b_0; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{array} \right.$$

Параметры (6 штук): $\alpha, \sigma^2, \ell_0, b_0, \beta, \phi$.

ETS(AAdN): прогнозируем

Прогноз числа свадеб по ETS(AAdN)



Прогноз на 1 шаг вперёд

$$\left\{ \begin{array}{l} y_t = \ell_{t-1} + \phi b_{t-1} + u_t; \\ \ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha u_t, \text{ стартовое } \ell_0; \\ b_t = \phi b_{t-1} + \beta u_t, \text{ стартовое } b_0; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{array} \right.$$

$$y_{T+1} = \ell_T + \phi b_T + u_{T+1}$$

$$(y_{T+1} \mid \mathcal{F}_T) \sim \mathcal{N}(\ell_T + \phi b_T; \sigma^2)$$

Прогноз на 2 шага вперёд

$$\begin{cases} y_t = \ell_{t-1} + \phi b_{t-1} + u_t; \\ \ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha u_t, \text{ стартовое } \ell_0; \\ b_t = \phi b_{t-1} + \beta u_t, \text{ стартовое } b_0; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{cases}$$

$$y_{T+2} = \ell_{T+1} + \phi b_{T+1} + u_{T+2} = (\ell_T + \phi b_T + \alpha u_{T+1}) + \\ + \phi(\phi b_T + \beta u_{T+1}) + u_{T+2}$$

$$(y_{T+2} \mid \mathcal{F}_T) \sim \mathcal{N}(\ell_T + (\phi + \phi^2)b_T; \sigma^2((\alpha + \phi\beta)^2 + 1))$$

ETS(AAdN): итоги

- На малом горизонте прогнозирования **тренд есть**.
- На большом горизонте прогнозирования **тренда нет**.
- **Один** дополнительный параметр.
- Можно получить ETS(AAdA) модель **с сезонностью**.

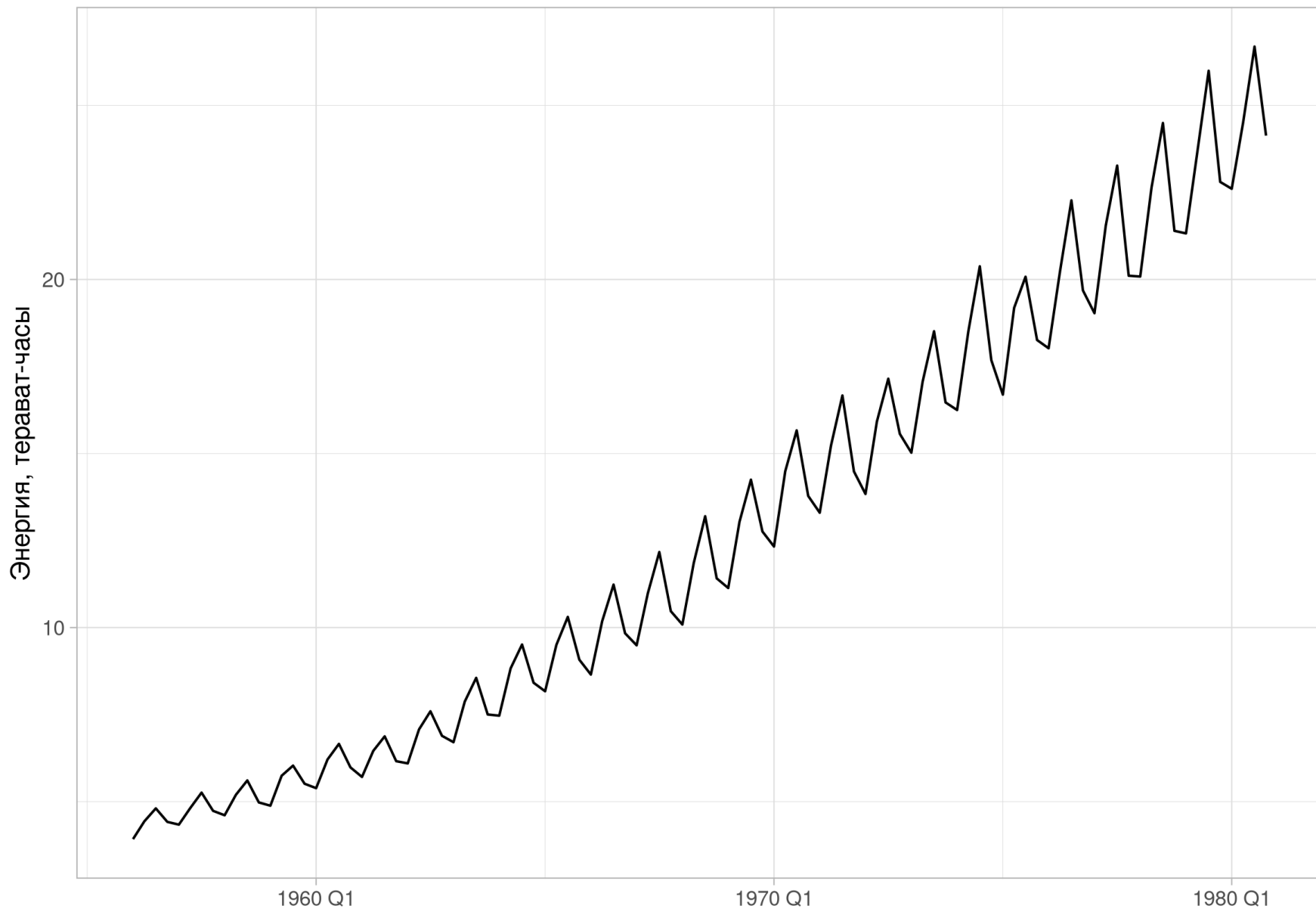
ETS: мультипликативные компоненты

ETS: мультипликативные компоненты

- Мультипликативные составляющие.
- Формулы для прогнозов.

Разная амплитуда колебаний

Производство электроэнергии в Австралии



Разная амплитуда колебаний

Возможные **решения**:

- Переход к логарифмам, $y_t \rightarrow \ln y_t$.
- Преобразование Бокса-Кокса, $y_t \rightarrow bc(y_t, \lambda)$.
- Мультипликативные компоненты.

ETS(MNM): уравнения

ETS(MNM) для месячных данных:

$$\begin{cases} y_t = \ell_{t-1} \cdot s_{t-12} \cdot (1 + u_t); \\ \ell_t = \ell_{t-1} \cdot (1 + \alpha u_t), \text{ стартовое } \ell_0; \\ s_t = s_{t-12} \cdot (1 + \gamma u_t), \text{ стартовые } s_0, \dots, s_{-11}; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{cases}$$

ETS(ANA):

$$\begin{cases} y_t = \ell_{t-1} + s_{t-12} + u_t; \\ \ell_t = \ell_{t-1} + \alpha u_t, \text{ стартовое } \ell_0; \\ s_t = s_{t-12} + \gamma u_t, \text{ стартовые } s_0, \dots, s_{-11}; \end{cases}$$

ETS(MNM): параметры

ETS(MNM) для месячных данных:

$$\left\{ \begin{array}{l} y_t = \ell_{t-1} \cdot s_{t-12} \cdot (1 + u_t); \\ \ell_t = \ell_{t-1} \cdot (1 + \alpha u_t), \text{ стартовое } \ell_0; \\ s_t = s_{t-12} \cdot (1 + \gamma u_t), \text{ стартовые } s_0, \dots, s_{-11}; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{array} \right.$$

Несезонные параметры: α, σ^2, ℓ_0 .

Сезонные параметры: $\gamma, s_0, s_{-1}, \dots, s_{-11}$.

Ограничение: $s_0 \cdot s_{-1} \cdot \dots \cdot s_{-11} = 1$.

Всего: 15 параметров.

Единицы измерения

Ряды y_t, ℓ_t — **исходные** единицы измерения.

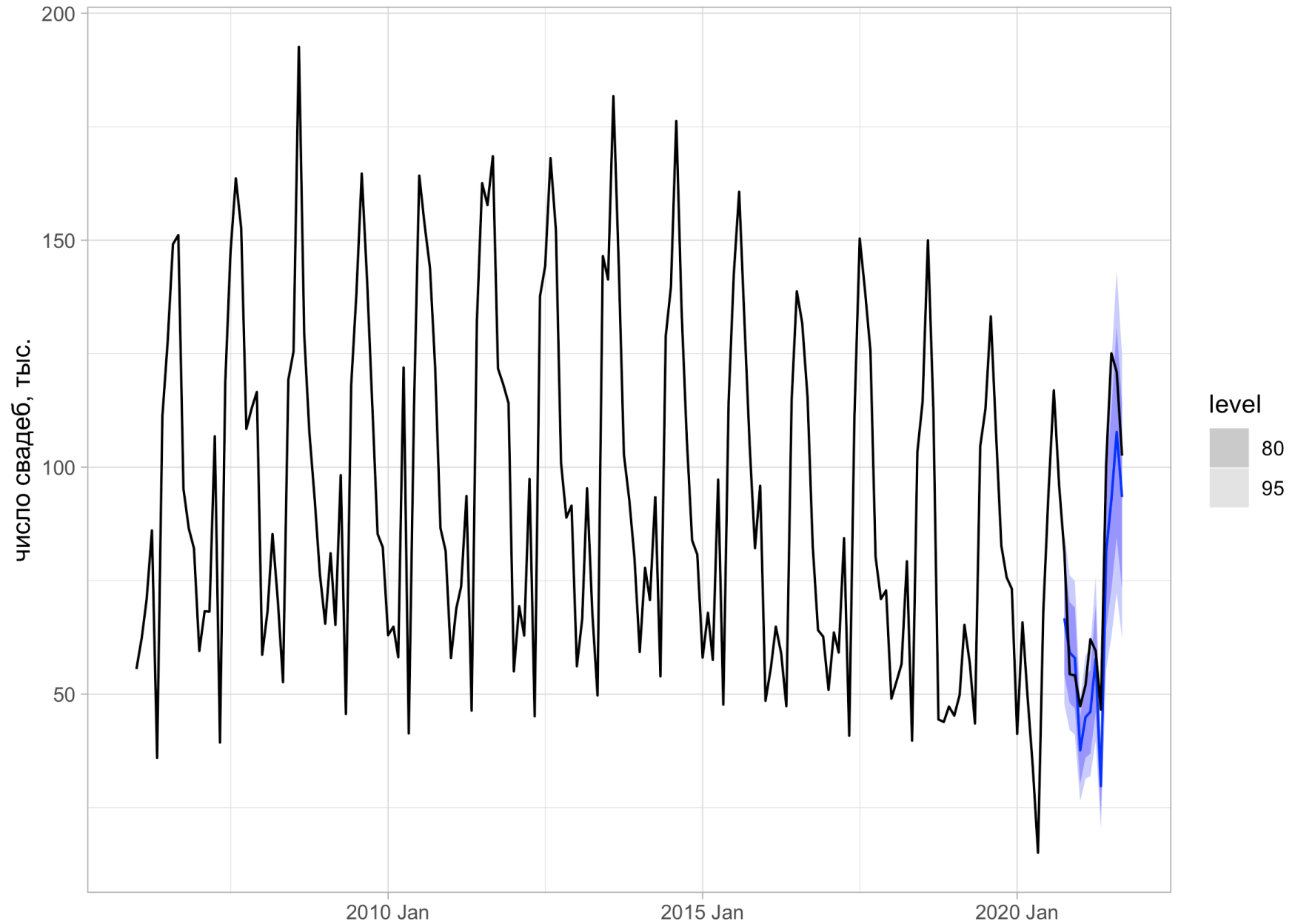
Ряды s_t, u_t — доли.

Ряд s_t измеряется относительно единицы, например, $s_t = 0.9$ — ниже тренда на 10%.

Ряд u_t измеряется относительно нуля, например, $u_t = -0.1$ — падение на 10%.

ETS(MNM): прогнозируем

Прогноз числа свадеб по ETS(MNM)



Прогноз на 1 шаг вперёд

$$\left\{ \begin{array}{l} y_t = \ell_{t-1} \cdot s_{t-12} \cdot (1 + u_t); \\ \ell_t = \ell_{t-1} \cdot (1 + \alpha u_t), \text{ стартовое } \ell_0; \\ s_t = s_{t-12} \cdot (1 + \gamma u_t), \text{ стартовые } s_0, \dots, s_{-11}; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{array} \right.$$

$$y_{T+1} = \ell_T \cdot s_{T-11} \cdot (1 + u_{T+1})$$

$$(y_{T+1} \mid \mathcal{F}_T) \sim \mathcal{N}(\ell_T \cdot s_{T-11}; (\ell_T \cdot s_{T-11})^2 \sigma^2)$$

Прогноз на 2 шага вперёд

$$\left\{ \begin{array}{l} y_t = \ell_{t-1} \cdot s_{t-12} \cdot (1 + u_t); \\ \ell_t = \ell_{t-1} \cdot (1 + \alpha u_t), \text{ стартовое } \ell_0; \\ s_t = s_{t-12} \cdot (1 + \gamma u_t), \text{ стартовые } s_0, \dots, s_{-11}; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{array} \right.$$

$$\begin{aligned} y_{T+2} &= \ell_{T+1} \cdot s_{T-10} \cdot (1 + u_{T+2}) = \\ &= \ell_T (1 + \alpha u_{T+1}) \cdot s_{T-10} \cdot (1 + u_{T+2}) \end{aligned}$$

$$(y_{T+2} \mid \mathcal{F}_T) \dot{\sim} \mathcal{N}(\ell_T \cdot s_{T-10}; \dots)$$

Мультипликативная ETS: итоги

- Моделирует **разную** амплитуду колебаний.
- Для **положительных рядов**.
- Простор для новых **комбинаций**.

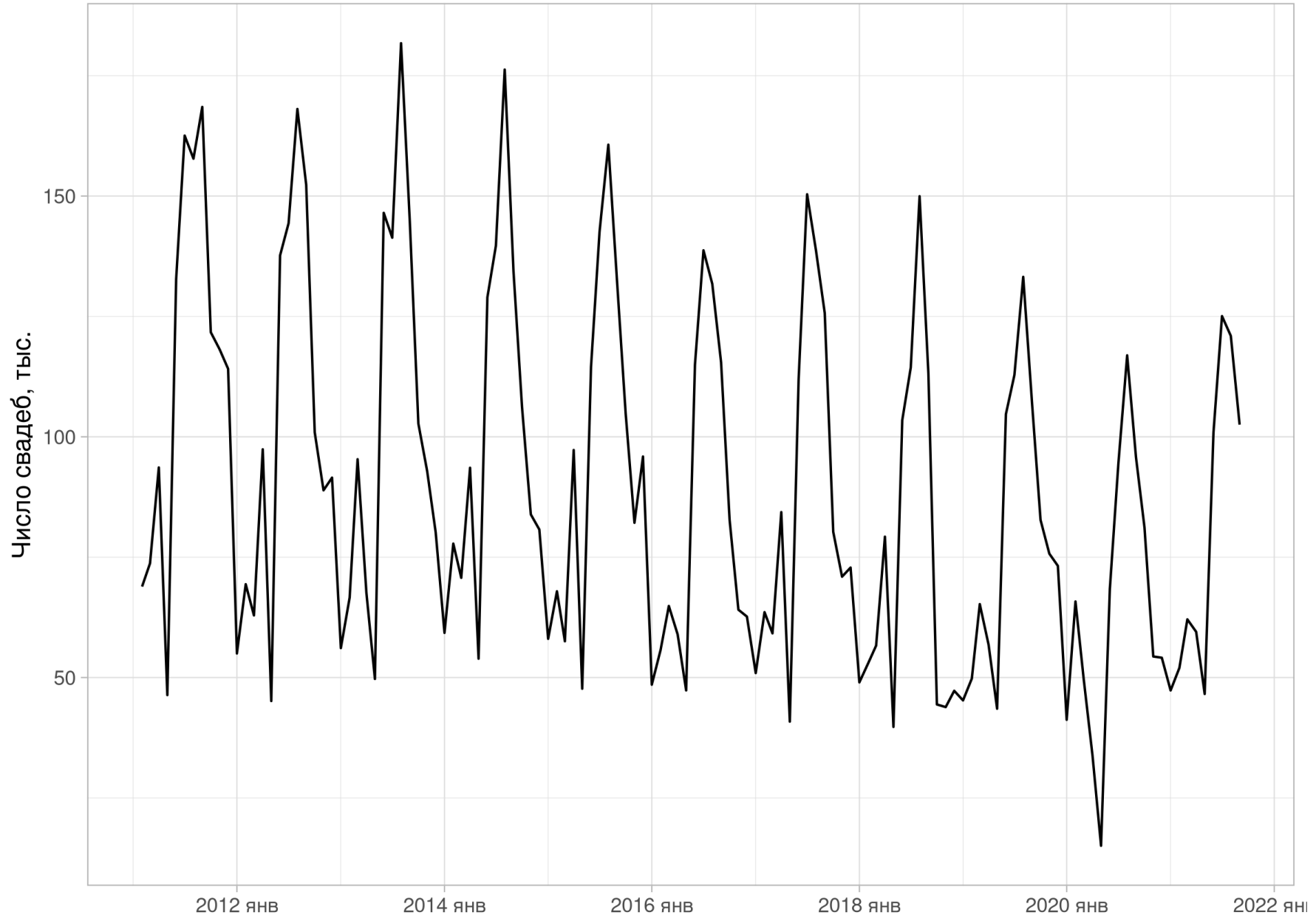
Собери свой ETS!

Собери свой ETS: план

- Собираем ETS(MAdM) модель.
- Прогнозы.

Разная амплитуда колебаний

Динамика числа свадеб в России



Хочу разные компоненты

Сезонность похожа на мультипликативную.

Мультипликативный тренд означал бы экспоненциальный рост.

Хочу аддитивный затухающий тренд.

ETS(MAdM): уравнения

ETS(MNM) для месячных данных:

$$\left\{ \begin{array}{l} y_t = \ell_{t-1} \cdot s_{t-12} \cdot (1 + u_t); \\ \ell_t = \ell_{t-1} \cdot (1 + \alpha u_t), \text{ стартовое } \ell_0; \\ s_t = s_{t-12} \cdot (1 + \gamma u_t), \text{ стартовые } s_0, \dots, s_{-11}; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{array} \right.$$

Как сюда добавить аддитивный тренд?

$$b_t = \phi b_{t-1} + \beta u_t, \text{ стартовое } b_0.$$

ETS(MAdM): уравнения

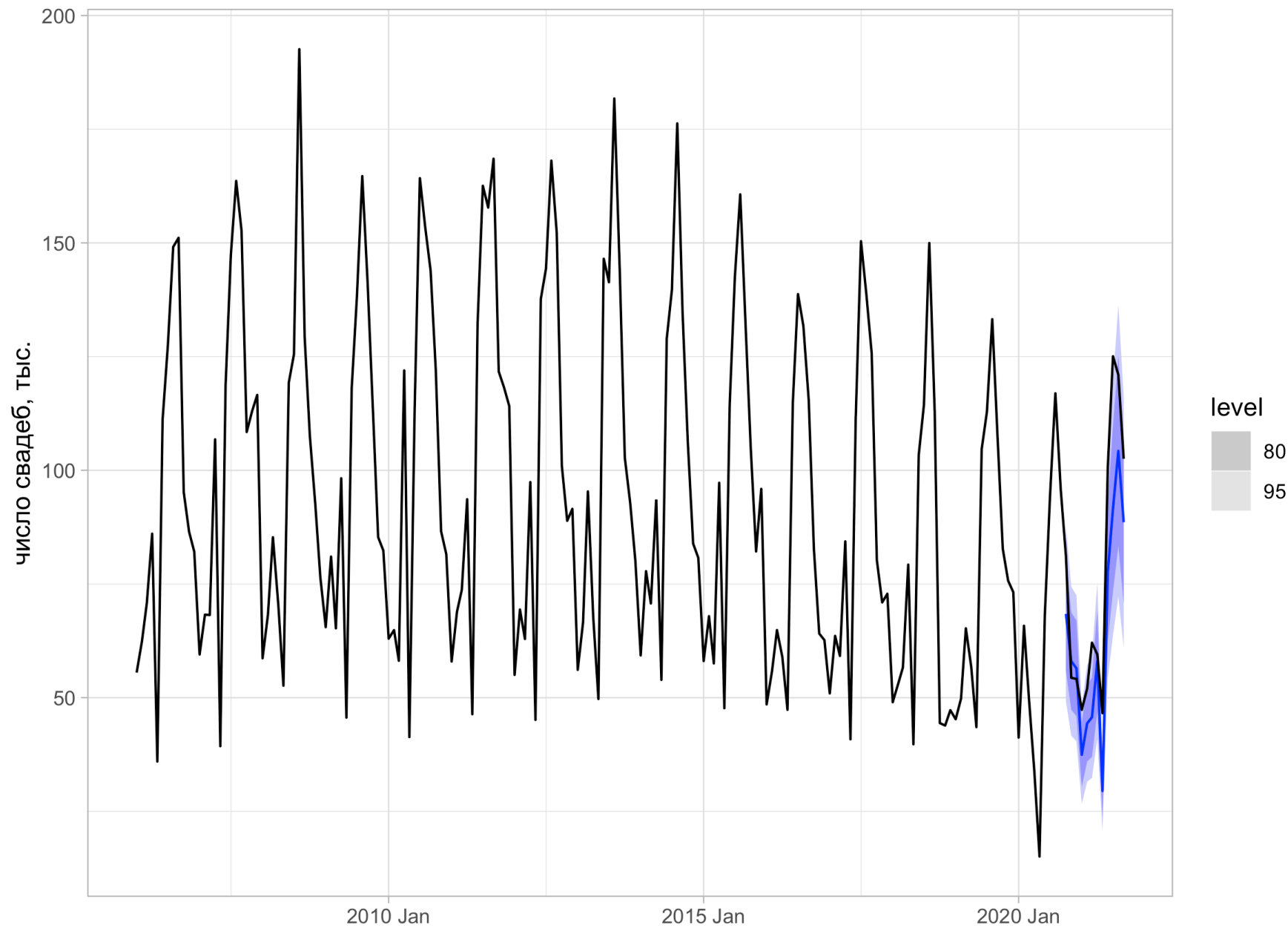
ETS(MAdM) для месячных данных:

$$\left\{ \begin{array}{l} y_t = (\ell_{t-1} + \phi b_{t-1}) \cdot s_{t-12} \cdot (1 + u_t); \\ \ell_t = (\ell_{t-1} + \phi b_{t-1}) \cdot (1 + \alpha u_t), \text{ стартовое } \ell_0; \\ b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})u_t, \text{ стартовое } b_0; \\ s_t = s_{t-12} \cdot (1 + \gamma u_t), \text{ стартовые } s_0, \dots, s_{-11}; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{array} \right.$$

Параметры — 18 штук.

ETS(MAdM): прогнозируем

Прогноз числа свадеб по ETS(MAdM)



Прогноз на 1 шаг вперёд

$$\left\{ \begin{array}{l} y_t = (\ell_{t-1} + \phi b_{t-1}) \cdot s_{t-12} \cdot (1 + u_t); \\ \ell_t = (\ell_{t-1} + \phi b_{t-1}) \cdot (1 + \alpha u_t), \text{ стартовое } \ell_0; \\ b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})u_t, \text{ стартовое } b_0; \\ s_t = s_{t-12} \cdot (1 + \gamma u_t), \text{ стартовые } s_0, \dots, s_{-11}; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{array} \right.$$

$$y_{T+1} = (\ell_T + \phi b_T) \cdot s_{T-11} \cdot (1 + u_{T+1})$$

$$(y_{T+1} \mid \mathcal{F}_T) \sim \mathcal{N}((\ell_T + \phi b_T) \cdot s_{T-11}; (\ell_T + \phi b_T)^2 \cdot s_{T-11}^2 \sigma^2)$$

Сколько всего ETS моделей?

Ошибка: A, M.

Тренд: N, A, Ad, M, Md.

Сезонность: N, A, M.

A — аддитивная составляющая.

M — мультипликативная составляющая.

N — нет составляющей.

d — дампирование для тренда.

Формально: 30 вариантов.

Исторические названия

ETS(ANN) — простое экспоненциальное сглаживание.

ETS(AAA) — аддитивный метод Хольта-Винтерса.

ETS(AAM) — мультипликативный метод Хольта-Винтерса.

ETS(AAdM) — метод Хольта-Винтерса с затухающим трендом.

Какой вариант выбрать?

Разная амплитуда колебаний: признак мультипликативных моделей.

Работает автоматический выбор по критерию AIC.

Часть мультипликативных моделей может быть численно неустойчива или не реализованы в софте.

Собери свой ETS: итоги

- Можно смешивать разные компоненты.
- **Ошибка**: A, M.
- **Тренд**: N, A, Ad, M, Md.
- **Сезонность**: N, A, M.
- Некоторые комбинации могут быть **неустойчивы**.

Тета-метод

Тета-метод: план

- Неожиданный лидер.
- Авторская версия.
- Частный случай ETS.

Тета-метод

Появился в 2000 году и стал сенсацией на соревнованиях МЗ по прогнозированию рядов.

Работает для несезонных рядов.

Изначально без статистической модели.

Авторская версия

1. Раскладываем ряд на две тета-линии ($\theta = 0, \theta = 2$).
2. Прогнозируем нулевую линию с помощью линейной регрессии.
3. Прогнозируем вторую линию с помощью ETS(ANN).
4. Усредняем прогнозы.

Можно предварительно удалить сезонность и в конце вернуть обратно.

Что такое тета-линия?

Нулевая тета-линия — регрессия ряда на время:

$$\hat{y}_t = \hat{\beta}_1 + \hat{\beta}_2 t.$$

Тета линия для произвольного тета:

$$\Delta^2 y_t^{new} = \theta \Delta^2 y_t.$$

Интуиция

- Нулевая тета-линия ловит долгосрочную тенденцию ряда.
- Тета-линия ($\theta = 2$) ловит краткосрочную тенденцию.
Ускорение тета-линии в θ раза сильнее ускорения исходного ряда.
- Усреднение снижает дисперсию прогнозов.

Как подбирается тета-линия?

Берём $\theta = 2$:

$$\Delta^2 y_t^{new} = 2\Delta^2 y_t.$$

Или

$$y_t^{new} - 2y_{t-1}^{new} + y_{t-2}^{new} = 2(y_t - 2y_{t-1} + y_{t-2}).$$

Новый ряд y_t^{new} полностью определяется y_1^{new}, y_2^{new} .

Решаем оптимизационную задачу:

$$\sum_{t=1}^T (y_t - y_t^{new})^2 \rightarrow \min.$$

Статистическая модель

Уже в 2003 году появилась модель:

$$\begin{cases} y_t = \ell_t + b + u_t; \\ \ell_t = \ell_{t-1} + b + \alpha u_t; \\ \ell_1 = y_1. \end{cases}$$

Или:

$$\Delta y_t = b + (\alpha - 1)u_{t-1} + u_t.$$

Тета-метод — вариант ETS

Основа — ETS(AAN):

$$\begin{cases} y_t = \ell_{t-1} + b_{t-1} + u_t; \\ \ell_t = \ell_{t-1} + b_{t-1} + \alpha u_t, \text{ стартовое } \ell_1; \\ b_t = b_{t-1} + \beta u_t, \text{ стартовое } b_0; \\ u_t \sim \mathcal{N}(0; \sigma^2) \text{ и независимы.} \end{cases}$$

Убираем стохастичность тренда $\beta = 0$.

Возможны **нюансы** инициализации.

Тета-метод: итоги

- Хорошо работает для **несезонных** данных.
- Особая **вариация** ETS модели.