Нелинейные модели регрессии

Виды нелинейных моделей регрессии

Нелинейная регрессия

Нелинейные регрессии по объясняющим переменным

- $\tilde{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2$;
- $\tilde{y} = \beta_0 + \beta_1 \dfrac{1}{x}$;

Нелинейные регрессии по оцениваемым параметрам

нелинейная модель внутренне линейна

- $\tilde{y} = \beta_0 \cdot x^{\beta_1}$;
- $\tilde{y} = e^{\beta_1 + \beta_2 \cdot x}$;
- $\tilde{y} = \beta_0 \cdot \beta_1^{\,x}$

нелинейная модель внутренне нелинейна

- $\tilde{y} = \beta_0 + \beta_1 x^{\beta_2}$;
- $\tilde{y} = \beta_0 \left(1 - \dfrac{1}{1 - x^{\beta_1}}\right)$

Подходы к оцениванию параметров нелинейных моделей регрессии

1. Линеаризация - подбор преобразований к анализируемым переменным $y, x_1, x_2, \ldots, x_k$, которые позволили бы представить искомую зависимость в виде линейного соотношения между преобразованными переменными; другими словами.

Если $\varphi_0, \varphi_1, \ldots \varphi_p$ - искомые функции, которые определяют переход к преобразованным переменным, т.е. $y^* = \varphi_0(y),\ \ x_1^* = \varphi_1(x_1),\ \ \ldots\ \ x_p^* = \varphi_p(x_p)$, то связь между y и $X = (x_1, x_2, \ldots, x_k)$ может быть представлена в виде линейной функции регрессии y^* от X^*, а именно:

$$y_i^* = \beta_0 + \beta_1 x_i^* + \ldots + \beta_p x_p^* + \varepsilon_i, \qquad i = 1, 2, \ldots, n$$

Подходы к оцениванию параметров нелинейных моделей регрессии

2. В случае **невозможности линеаризации модели** исследуется искомая регрессионная зависимость в терминах исходных переменных:

$$y_i = f\left(X_i, \beta\right) + \varepsilon_i .$$

Если спецификация регрессионных остатков ε_i соответствует условиям классической модели, то для вычисления МНК-оценок решается оптимизационная задача вида:

$$b_{МНК} = \arg \min_b \sum_{i=1}^{n} (y_i - f(X_i, \beta))^2$$

Некоторые виды нелинейных зависимостей, поддающиеся линеаризации

Зависимость гиперболического типа

1. $\acute{o}_i = \beta_0 + \beta_1 \frac{1}{x_i} + \varepsilon_i \quad (0 < x_i < \infty)$

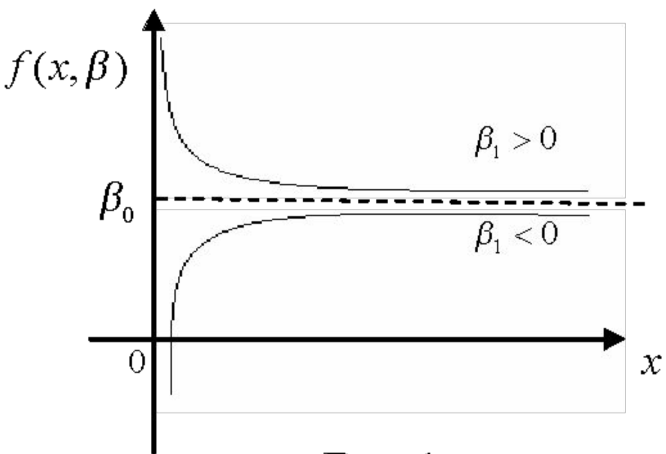


Рис.1

Преобразования объясняющей переменной $x^* = \frac{1}{x}$, $X^* = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \frac{1}{x_1} & \frac{1}{x_2} & \dots & \frac{1}{x_n} \end{pmatrix}^T$

$$\acute{o}_i = \beta_0 + \beta_1 x_i^* + \varepsilon_i \,.$$

$$\textbf{2.}\quad y_i = \frac{1}{\beta_0 + \beta_1 x_i + \varepsilon_i},\ (-\frac{\beta_0}{\beta_1} < x < \infty)$$

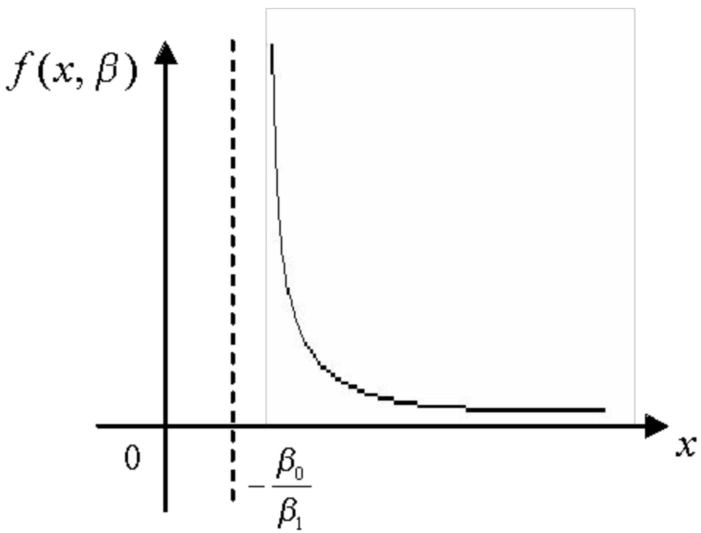


Рис. 2 *а*) $\beta_0 < 0,\quad \beta_1 > 0$

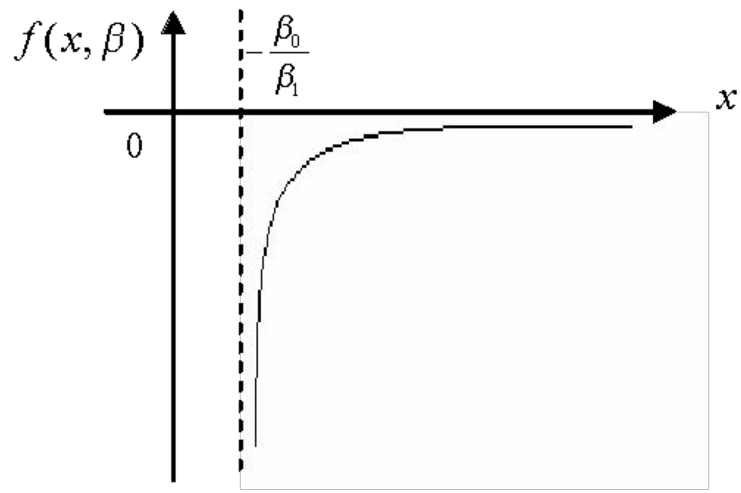


Рис. 2 *б*) $\beta_0 > 0,\quad \beta_1 < 0$

Преобразования результативного признака $y^* = \frac{1}{y},\quad Y^* = \left(\frac{1}{y_1}\quad \frac{1}{y_2}\quad \dots\quad \frac{1}{y_n}\right)^T$

$$\acute{o}^{*}{}_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Показательная (экспоненциальная) зависимость

$$\mathbf{1} \quad y_i = \beta_0 e^{\beta_1 x+\varepsilon}$$

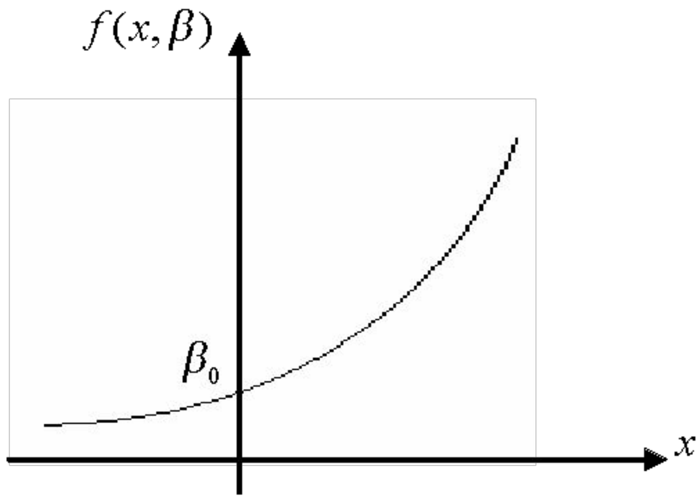


Рис. 1 *а*) $\beta_1 > 0$

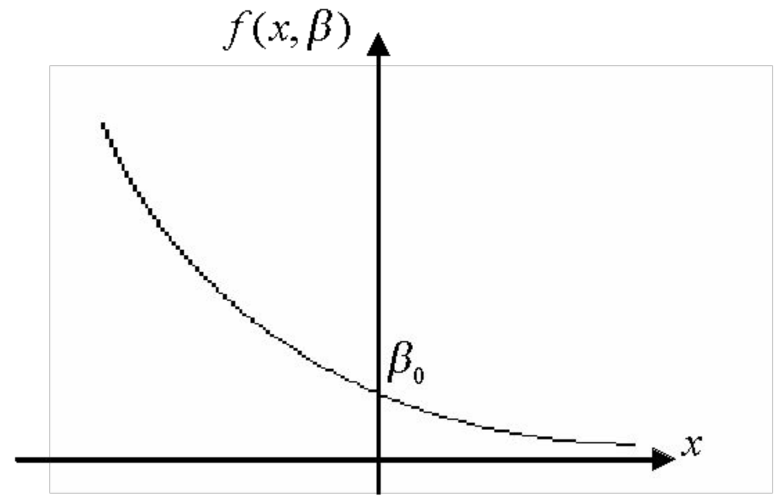


Рис. 1 *б*) $\beta_1 < 0$

Преобразования результативного признака $y^* = \ln y$,

$Y^* = (\ln y_1 \quad \ln y_2 \quad \dots \quad \ln y_n)^T$

$$y^*_i = \beta_0^* + \beta_1 \tilde{o}_i + \varepsilon_i \text{ , } \beta_0^* = \ln \beta_0$$

2 $y_i = \beta_0 e^{\frac{\beta_1}{x_i}+\varepsilon_i}$

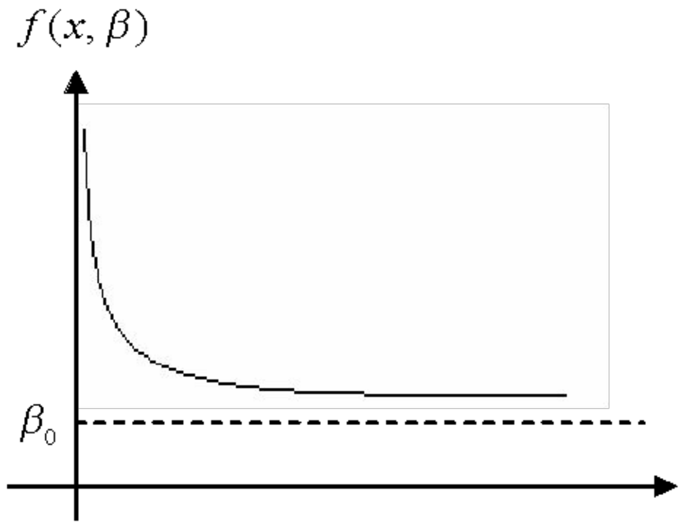


Рис. 2 *а*) $\beta_1 > 0$

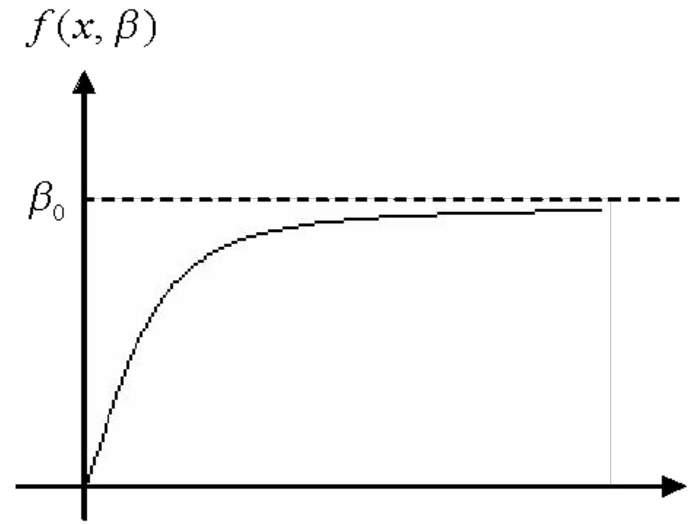
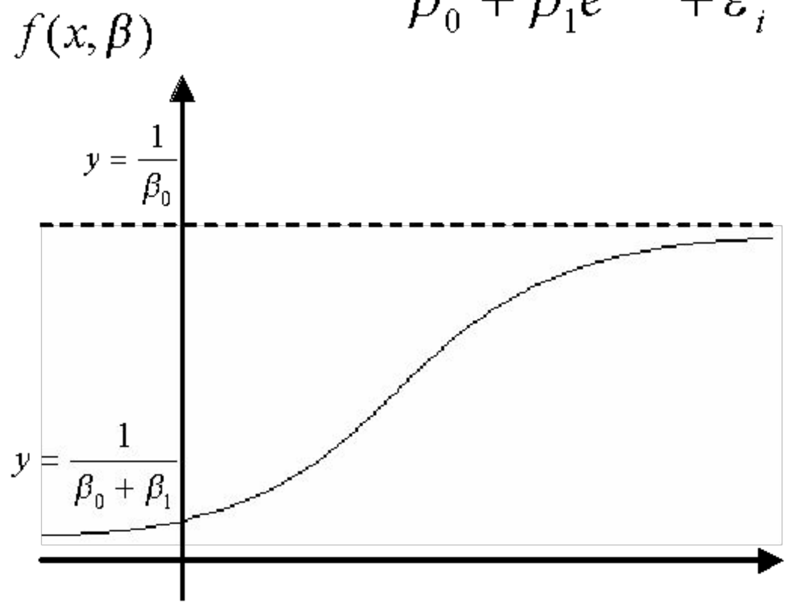


Рис. 2 *б*) $\beta_1 < 0$

Преобразования переменных $y^* = \ln y,\ x^* = \dfrac{1}{x}$. Где $\beta_0^* = \ln \beta_0$

$Y^* = \left(\ln y_1 \quad \ln y_2 \quad \ldots \quad \ln y_n\right)^T$ и матрица $X^* = \begin{pmatrix} 1 & 1 & \ldots & 1 \\ \dfrac{1}{x_1} & \dfrac{1}{x_2} & \ldots. & \dfrac{1}{x_n} \end{pmatrix}^T$

3 Логистическая кривая $y_i = \dfrac{1}{\beta_0 + \beta_1 e^{-x_i} + \varepsilon_i}, \; 0 \le x < \infty$



$\beta_1 > 0$

Преобразования переменных $y^* = \dfrac{1}{y}, \quad x^* = e^{-x}, \quad Y^* = \begin{pmatrix} \dfrac{1}{y_1} & \dfrac{1}{y_2} & \dots & \dfrac{1}{y_n} \end{pmatrix}^T$

$$X^* = \begin{pmatrix} 1 & 1 & \dots & 1 \\ e^{-x_1} & e^{-x_2} & \dots. & e^{-x_n} \end{pmatrix}^T$$

Зависимость степенного типа

$$y_i = \beta_0 x_i^{\beta_1}$$

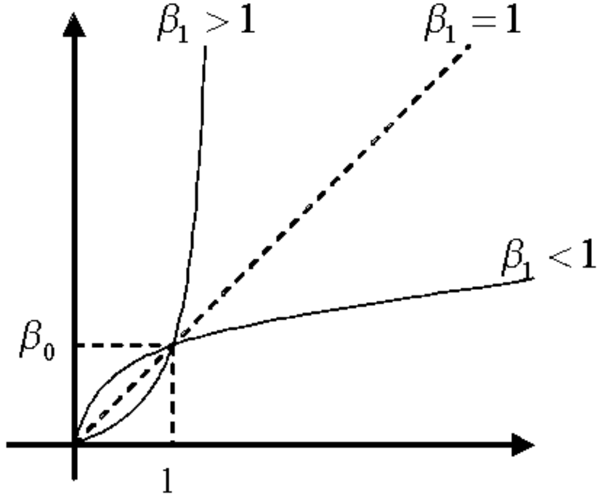


Рис. 1 *а*) $\beta_1 > 0$

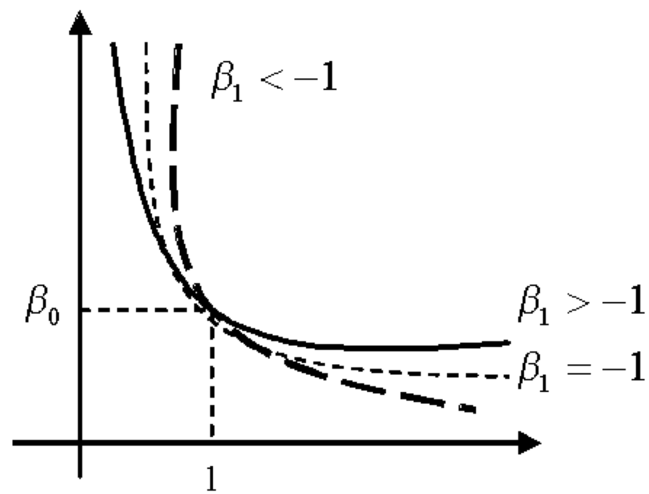


Рис. 1 *б*) $\beta_1 < 0$

Преобразования переменных $y^* = \ln y$, $x^* = \ln x$, где $\beta_0^* = \ln \beta_0$

Важную роль зависимости степенного типа играют в задачах построения и анализа производственных функций, функций спроса. При анализе степенных регрессионных зависимостей содержательную интерпретацию получает коэффициент β_1 как коэффициент эластичности.

Метод Бокса-Кокса

Метод основан на степенном преобразовании переменных:

$$Y(\lambda)=\frac{Y^{\lambda}-1}{\lambda},\ X(\lambda)=\frac{X^{\lambda}-1}{\lambda},\ \lambda\neq 0.$$

Линейная регрессия с учетом степенной трансформации переменных примет вид:

$$Y(\lambda)=\beta_0+\beta_1 X_1(\lambda)+\beta_2 X_2(\lambda)+\ldots+\beta_k X_k(\lambda)+\varepsilon.$$

Задача состоит в определении оптимального λ. Наилучшим считается то значение λ, при котором достигается максимум логарифма функции правдоподобия:

$$\ln L(\lambda)=-\frac{n}{2}\ln\left[\sigma^2(\lambda)\right]+(\lambda-1)\sum_{i=1}^{n}\ln y_i\ ,$$

где $\sigma^2(\lambda)$ – это оценка наибольшего правдоподобия для σ^2 при данном λ.

Значение λ подбирают из диапазона от -2 до +2.