MODELE NIELINIOWE

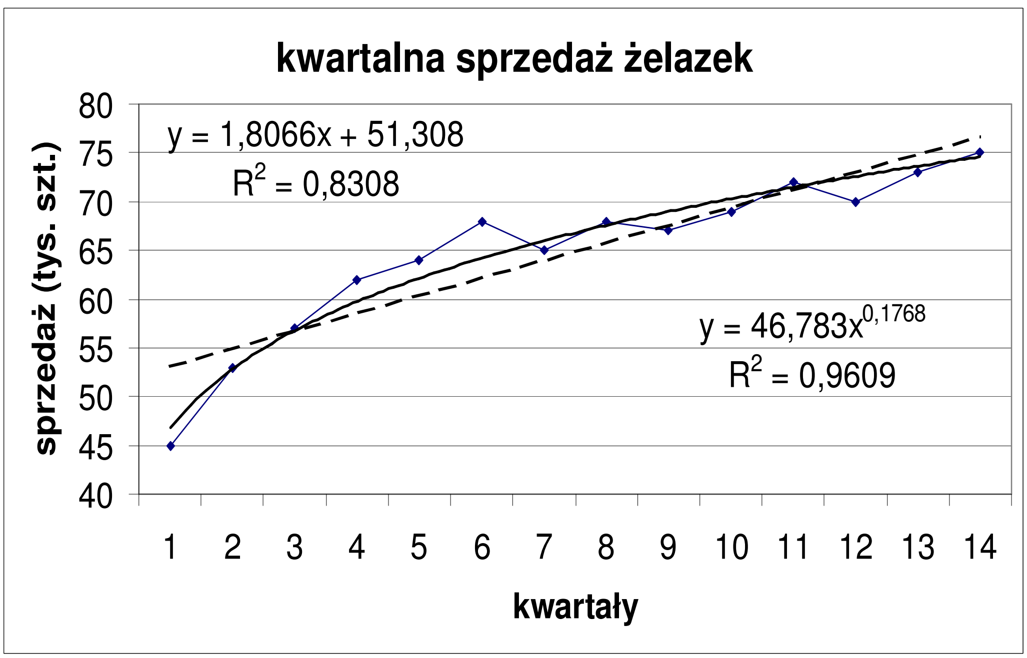
Przykład. (model liniowy i nieliniowy - porównanie).
Kwartalna sprzedaż żelazek w latach 1996 - 1999 wynosiła:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14
y_t	45	53	57	62	64	68	65	68	67	69	72	70	73	75

Wyznacz prognozę na trzeci kwartał roku 1999.

Z wykresu widać, że trend jest rosnący lecz coraz wolniejszy (względne nasycenie rynku) zatem zamiast trendu liniowego należy zastosować trend nieliniowy o malejącym tempie wzrostu np. model potęgowy $y = at^b$ $(0 < b < 1)$.

Logarytmując model potęgowy $y = at^b$ otrzymamy $\ln y = \ln a + b\ln t$ a po podstawieniu $t' = \ln t$; $y' = \ln y$; $a' = \ln a$ otrzymamy model liniowy $y' = a' + b't'$.
Wykres par (lnt, lny) potwierdza trafność wybranego modelu.



t	1	2	3	4	5	6	7	8	9	10	11	12	13	14
lnt	0	0,69	1,1	1,39	1,61	1,79	1,95	2,08	2,2	2,3	2,4	2,48	2,56	2,64
lny	3,81	3,97	4,04	4,13	4,16	4,22	4,17	4,22	4,2	4,23	4,28	4,25	4,29	4,32

Oszacowany model liniowy ma postać:

$$\hat{y}' = 3{,}846 + 0{,}177t'$$

Oszacowany model potęgowy ma postać:

$$\hat{y} = 46{,}8t^{0{,}177}$$

Zatem prognoza punktowa wyniesie $y^*_{15} = 46{,}8 \cdot 15^{0{,}177} = 75{,}6$

Problem: wyznaczyć prognozę punktową na 16 kwartał.

Model liniowy względem parametrów.

$$\hat{Y} = b_0 + b_1 f_1(X_1) + ... + b_k f_k(X_k)$$

$f_1, ..., f_k$ - funkcje różnowartościowe.

Podstawiając $Z_i = f_i(X_i)$ $i = 1, ..., k$ otrzymamy model liniowy:

$$\hat{Y} = b_0 + b_1 Z_1 + ... + b_k Z_k$$

Przykład. (model paraboliczny).

$$\hat{Y} = b_0 + b_1 X + b_2 X^2 \qquad b_2 \neq 0,$$

podstawiając

$Z_1 = X$, $Z_2 = X^2$, otrzymamy model liniowy

$$\hat{Y} = b_0 + b_1 Z_1 + b_2 Z_2$$

Przykład. (model hiperboliczny).

$$\hat{Y} = b_0 + \frac{b_1}{X} \qquad X > 0 \qquad b_1 \neq 0,$$

podstawiając

$Z = 1/X$, otrzymamy model liniowy

$$\hat{Y} = b_0 + b_1 Z_1$$

Przykład. (model logarytmiczny).

$$\hat{Y} = b_0 + b_1 \ln X \qquad X > 0 \qquad b_1 \neq 0,$$

podstawiając

$Z = \ln X$, otrzymamy model liniowy

$$\hat{Y} = b_0 + b_1 Z_1$$

Model linearyzowany.

Model linearyzowany to taki model dla którego istnieje jednoznaczne przekształcenie obu jego stron takie, że otrzymamy model liniowy lub liniowy względem parametrów.

Przykład. (model wykładniczy).

$$\hat{Y} = b_0 p^{b_1 X} \qquad b_0,\ p > 0,\ \ p \neq 1$$

ogólnie

$$\hat{Y} = b_0 p^{b_1 X_1 + \ldots + b_k X_k}$$

Linearyzacja polega na obustronnym zlogarytmowaniu modelu wyjściowego (najlepiej zastosować logarytm o podstawie p) i podstawieniu

$$\tilde{b}_0 = \log_p b_0, \qquad \hat{Z} = \log_p \hat{Y},$$

otrzymamy model liniowy

$$\hat{Z} = \tilde{b}_0 + b_1 X_1 + \ldots + b_k X_k$$

Przykład. (model potęgowy).

$$\hat{Y} = b_0 X^{b_1} \qquad b_0,\ X > 0,$$

ogólnie

$$\hat{Y} = b_0 X_1^{b_1} X_2^{b_2} \ldots X_k^{b_k}$$

Linearyzacja polega na obustronnym zlogarytmowaniu modelu wyjściowego (najlepiej zastosować logarytm naturalny - o podstawie e) i podstawieniu

$$\tilde{b}_0 = \ln b_0, \quad Z_i = \ln X_i, \qquad \hat{Z} = \ln \hat{Y},$$

otrzymamy model liniowy

$$\hat{Z} = \tilde{b}_0 + b_1 Z_1 + \ldots + b_k Z_k$$

Estymacja parametrów funkcji regresji krzywoliniowej.

Parametry wybranej funkcji nieliniowej wyznacza się też bezpośrednio metodą najmniejszych kwadratów, korzystając z odpowiedniego układu równań normalnych.

Funkcja wielomianowa.

$$\hat{y}_i = b_0 + b_1 x + b_2 x^2 + \ldots + b_k x^k$$

Jej parametry $b_0, \ldots, b_k$ wyznaczamy rozwiązując układ równań normalnych który ma postać:

$$\begin{cases} nb_0 + \left(\sum_i x_i\right) b_1 + \left(\sum_i x_i^2\right) b_2 + \ldots + \left(\sum_i x_i^k\right) b_k = \sum_i y_i \\ \left(\sum_i x_i\right) b_0 + \left(\sum_i x_i^2\right) b_1 + \left(\sum_i x_i^3\right) b_2 + \ldots + \left(\sum_i x_i^{k+1}\right) b_k = \sum_i x_i y_i \\ \ldots \\ \left(\sum_i x_i^k\right) b_0 + \left(\sum_i x_i^{k+1}\right) b_1 + \left(\sum_i x_i^{k+2}\right) b_2 + \ldots + \left(\sum_i x_i^{2k}\right) b_k = \sum_i x_i^k y_i \end{cases}$$

Powyższy układ k + 1 równań otrzymujemy przyrównując do zera pochodne cząstkowe funkcji k + 1 zmiennych

$$S(b_0, b_1, b_2, \ldots, b_k) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} \left(y_i - \left(b_0 + b_1 x_i + b_2 x_i^2 + \ldots + b_k x_i^k\right)\right)^2$$

W szczególności **funkcja kwadratowa.**

$$\hat{y}_i = b_0 + b_1 x + b_2 x^2$$

Jej parametry b_0, b_1, b_2 wyznaczamy rozwiązując układ równań normalnych który ma postać:

$$\begin{cases} nb_0 + \left(\sum_i x_i\right) b_1 + \left(\sum_i x_i^2\right) b_2 = \sum_i y_i \\ \left(\sum_i x_i\right) b_0 + \left(\sum_i x_i^2\right) b_1 + \left(\sum_i x_i^3\right) b_2 = \sum_i x_i y_i \\ \left(\sum_i x_i^2\right) b_0 + \left(\sum_i x_i^3\right) b_1 + \left(\sum_i x_i^4\right) b_2 = \sum_i x_i^2 y_i \end{cases}$$

Funkcja potęgowa.

$$\hat{y}_i = a x_i^{\,b}$$

Założenie $a > 0$, $b > 0$, $x_i > 0$.

Chociaż jest to niekiedy szczególny przypadek funkcji wielomianowej to warto rozpatrywać go również oddzielnie. Jej parametry a, b wyznaczamy przez przekształcenie do postaci liniowej (logarytmujemy obie strony).

$$\ln \hat{y}_i = \ln a + b \ln x_i$$

Układ równań normalnych ma postać:

$$\begin{cases} \sum_i \ln y_i = n \ln a + b \sum_i \ln x_i \\ \sum_i \ln x_i \ln y_i = \ln a \sum_i \ln x_i + b \sum_i \ln^2 x_i \end{cases}$$

Rozwiązując powyższy układ równań (liniowych względem a' = lna i b) obliczamy a' i b.

Stąd $a = e^{a'}$.

Parametr b jest interpretowany jako współczynnik elastyczności, tzn. jeśli zmienna X wzrośnie o 1%, to Y zmieni się średnio o b %.

Funkcja wykładnicza.

$$\hat{y}_i = ab^{x_i}$$

Założenie a > 0, b > 0.
Logarytmując obie strony otrzymamy.

$$\ln \hat{y}_i = \ln a + x_i \ln b$$

Układ równań normalnych ma postać:

$$\begin{cases} \sum_i \ln y_i = n \ln a + \ln b \sum_i x_i \\ \sum_i x_i \ln y_i = \ln a \sum_i x_i + \ln b \sum_i x_i^2 \end{cases}$$

Rozwiązując powyższy układ równań (liniowych względem a' = lna i b' = lnb) obliczamy a' i b'. Stąd $a = e^{a'}$ i $b = e^{b'}$.

Parametr b jest interpretowany jako średni przyrost względny, tzn. jeśli zmienna X wzrośnie o jednostkę, to Y zmieni się średnio o (b - 1)100 %.

Funkcja logistyczna.

$$\hat{y}_i = \frac{a}{1 + be^{-ct}}$$

gdzie t - czas, a > 0, b > 1, c > 0.
Funkcja logistyczna służy między innymi do opisu i prognozowania wielkości sprzedaży produktu wchodzącego na rynek.
Przyjmujemy

$$z_t = \frac{1}{y_t}$$

Najpierw wyznaczamy wartości parametrów a, c

$$a = \frac{u}{u_0} - 1 \qquad\qquad c = -\ln \frac{u_1}{u}$$

gdzie

$$u = (n-1)\sum_{t=1}^{n-1} z_t^2 - \left(\sum_{t=1}^{n-1} z_t \right)^2$$

$$u_0 = \sum_{t=1}^{n-1} z_{t+1} \sum_{t=1}^{n-1} z_t^2 - \sum_{t=1}^{n-1} z_t \sum_{t=1}^{n-1} z_t \, z_{t+1}$$

$$u_1 = (n-1)\sum_{t=1}^{n-1} z_t \, z_{t+1} - \sum_{t=1}^{n-1} z_{t+1} \sum_{t=1}^{n-1} z_t$$

Następnie korzystając z obliczonych a i c obliczamy b

$$b = \frac{1}{n} \sum_{t=1}^{n} \left(\frac{a}{y_t} - 1 \right) e^{ct} \qquad\qquad t = 1, 2, ..., n$$

Parametr b > 1 gwarantuje istnienie punktu przegięcia, zaś parametr a jest interpretowany jako poziom nasycenia (asymptota pozioma).

Uwaga.

a) W przypadku nieliniowym miarą dopasowania modelu do danych statystycznych jest współczynnik korelacji krzywoliniowej

$$R=\sqrt{1-\frac{\sum_{i=1}^{n}(y_i-\hat{y})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}}; \qquad R \in <0,\ 1>.$$

b) Błąd prognozy dla modelu nieliniowego przekształconego do modelu liniowego za pomocą przekształcenia y' = f(y) gdzie f - funkcja różniczkowalna wyraża się wzorem

$$S_\tau = \frac{S_e\sqrt{1+\frac{1}{n}+\frac{(x_\tau-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}}{\left|\frac{dy'}{dy}\Big|y_\tau^*\right|}$$

W szczególności gdy y' = lny to $\frac{dy'}{dy}\Big|y_\tau^* = \frac{1}{y_\tau^*}$

y' = 1/y to $\frac{dy'}{dy}\Big|y_\tau^* = \frac{-1}{(y_\tau^*)^2}$

y' = e^y to $\frac{dy'}{dy}\Big|y_\tau^* = e^{y_\tau^*}$

c)

Przykład.

Mając dane

t	1	2	3	4	5	6	7	8
y	8	6	5	2	2	4	7	8

Wyznaczymy kwadratową funkcję regresji.

Korzystając z sum w poniższej tabeli układamy układ równań normalnych:

	t	y	t^2	t^3	t^4	y^2	ty	t^2y
	1	8	1	1	1	64	8	8
	2	6	4	8	16	36	12	24
	3	5	9	27	81	25	15	45
	4	2	16	64	256	4	8	32
	5	2	25	125	625	4	10	50
	6	4	36	216	1296	16	24	144
	7	7	49	343	2401	49	49	343
	8	8	64	512	4096	64	64	512
suma	36	42	204	1296	8772	262	190	1158

$a \cdot 8772 + b \cdot 1296 + c \cdot 204 = 1158$

$a \cdot 1296 + b \cdot 204 + c \cdot 36 = 190$

$a \cdot 204 + b \cdot 36 + c \cdot 8 = 42$

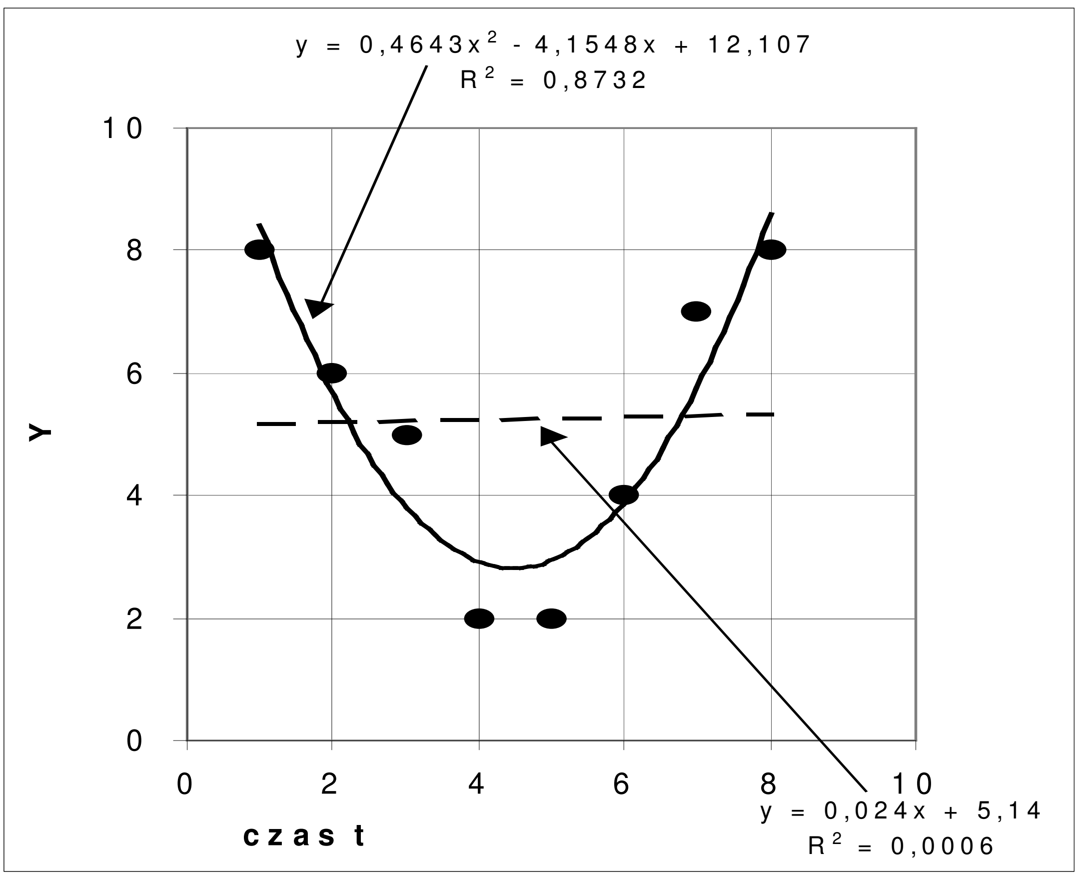
Rozwiązaniem (przybliżonym) układu jest
$a = 0{,}4643$; $b = -4{,}1548$; $c = 12{,}107$

Zatem funkcja regresji kwadratowej ma postać

$$\hat{y} = 0{,}4643x^2 - 4{,}1548x + 12{,}107$$

Funkcja ta jest dobrze dopasowana do danych statystycznych
($R^2 = 0{,}8732$, $r = 0{,}93$).

Zauważmy, że w tym przypadku funkcja liniowa nie jest dobrą funkcją regresji a bardzo niska wartość współczynnika korelacji Pearsona świadczy o braku zależności liniowej a nie o braku zależności jakiejkolwiek.



L. Kowalski, 18.03.2005