XI CONGRESSO BRASILEIRO DE ENGENHARIA DE AVALIAÇÕES E PERÍCIAS - XI COBREAP

TRANSFORMAÇÃO DE VARIÁVEIS - REGRESSÃO NÃO LINEAR

Giannakos, Isabela Beck da Silva
Engenheira Civil, CREA-RS 51967
IBAPE-RS 1025/95
Av. Cristóvão Colombo, 2834 cj. 603
Porto Alegre, RS – 90560-002
Tel. 0-XX-51 33379288 Fax 0-XX-51 33378977
***e-mail* bsgisa@terra.com.br - home page: www.bsg.com.br**

Leão, Manoel Luiz
Engenheiro Civil, CREA-RS 7125
IBAPE-RS 1020/97
Rua Miguel Tostes, 270
Porto Alegre, RS – 90430-060
Tel. 0-XX-51 33315618 Fax 0-XX-51 33325833
***e-mail* leao@orion.ufrgs.br**

Resumo:

A transformação da variável dependente, nos modelos de regressão, com o propósito de incrementar a linearidade dos mesmos, gera, ao desfazer-se a transformação, para direcionar a avaliação ao seu objeto final, uma função não-linear onde os parâmetros da função linear previamente definida não mais satisfazem o requisito do método de mínimos quadrados, cabendo buscar novos valores para os mesmos, a fim de preservar esta adesão ao requisito básico do citado método. O emprego, por outro lado, do valor unitário como variável dependente, cria situação análoga, com idêntica implicação.

Summary:

Real-estate appraisers frequently effect the transformatiom of the dependent variable in linear regression models, with the intent to improve their "goodness-of-fit". Later, when the original condition of the variable is replaced, to address the procedure's final aim, the result is actually obtained through a non-linear function,

while the intercept and slope of the preliminary linear function are maintained, although they no more attain que basic minimum-square requirement. Thus, a non-linear search for new parameters is required. The same occurs when the unit value (value per square meter) is chosen as the dependendent variable.

Introdução

Constitui prática generalizada entre avaliadores a de promover transformações nas variáveis consideradas no modelo de regressão adotado, atingindo tanto as variáveis independentes como as dependentes, com o propósito de, através desta mudança de escalas, obter comportamento mais linear para os dados da amostra e, assim, melhorar a representatividade e o poder estimador da equação de regressão, melhorias estas medidas pelo crescimento do coeficiente de determinação R^2.

Trabalho recentemente divulgado (Ref. 1), aponta para alguns problemas importantes, nem sempre devidamente atentados pelos avaliadores, que podem surgir nos procedimentos estatísticos de avaliação, quando transformada a variável dependente. Quando alterada apenas a variável independente, a função resultante continua linear e prevalecem os critérios de ajustamento por mínimos quadrados, que presidem todo o procedimento da regressão. Tal não ocorre, no entanto, quando modificada a variável dependente. O objetivo do presente trabalho é o de prosseguir na investigação deste tema.

A transformação da variável dependente e a "migração" de parâmetros

A transformação da variável dependente gera depois, quando desfeita a transformação, uma função não-linear para a determinação do valor do bem que constitui o objeto real da avaliação. Os parâmetros da equação linear, obtidos pelo procedimento de regressão, são utilizados para estimar o valor da variável dependente transformada. A seguir, no entanto, ao expressar o valor real do bem sob avaliação – e não mais o de seu inverso, de seu logaritmo, ou de qualquer outra transformação que haja o mesmo sofrido - o que se estará praticando, na realidade, é a utilização de uma função não-linear, com o "empréstimo" de parâmetros oriundos de uma função linear, distinta daquela, sem investigar se ditos parâmetros satisfazem a imposição básica, de oferecerem uma estimativa de "**mínimos quadrados**", isto é, a de provirem de uma equação que percorre o espaço da amostra colhida pelo perito, de forma a assegurar que seja mínimo o erro provável da estimativa, medido pela soma dos quadrados das diferenças entre os valores da variável dependente e as ordenadas da função estimadora, para as mesmas abscissas.

Mínimos quadrados – a determinação dos parâmetros

Seja um conjunto de "n" pares de valores X (variável independente) e Y (variável dependente), formando uma amostra, por exemplo, de dados de mercado, trazendo, para cada componente, uma característica X e um atributo Y, tais como área e preço. Calcular a equação de regressão linear que permita, dado X, estimar Y, corresponde a especificar os valores de "a" e "b", parâmetros (respectivamente linear e angular) da reta $Y_c = a + bX$, em que X é o valor da abscissa que caracteriza o bem objeto da avaliação e Y_c ("Y calculado") o respectivo valor estimado.

A propriedade fundamental desta reta é a de fazer com que seja a mínima possível a soma dos quadrados dos afastamentos $Y – Y_c$, em que Y_c representa as ordenadas da reta, para os valores de X associados a cada um dos valores de Y. Para obter os valores de a e b que satisfaçam esta imposição, extraem-se as derivadas parciais desta soma de quadrados, igualando-as, a seguir, a zero.

Seja $Z = S(Y-Y_c)^2$, lembrando que $Y_c = a + bX$.

Trata-se de minimizar Z; para tanto, tomam-se as derivadas parciais de Z em relação a, respectivamente, a e b, dZ/da e dZ/db, fazendo ambas iguais a zero. Resultam, daí, as equações, ditas "**normais**", a seguir:

$$SY = Na + bSX$$
$$SXY = aSX + bSX^2$$

Suponha-se, agora, que uma transformação de variáveis foi adotada, substituindo-se, por exemplo, Y por seu inverso $(1/Y_c)$, mantendo-se X na escala original. A reta de estimativa será, então, dada por

$$(1/Y_c) = a + bX; \qquad \text{(I)}$$

os parâmetros a e b, por sua vez, surgirão da solução do sistema de equações

$$S(1/Y) = Na + bSX \qquad \text{(II)}$$
$$S(X/Y) = aX + bSX^2 .$$

Há, pelo menos, duas maneiras para encontrar a solução:

a) Inverter a matriz dos coeficientes de a e b, acima (equações II), e multiplicar, a seguir, por esta matriz inversa, o vetor formado pelos termos independentes, S(1/Y) e S(X/Y). Resultarão os valores de a e b;

b) Recorrer ao comando "Regressão", oferecido pelas planilhas eletrônicas hoje disponíveis.

Isto posto, dado um valor de X, poder-se-á, então, estimar o correspondente valor de $1/Y_c$, o inverso de Y_c, pela equação I, acima. Mas, em se tratando, por exemplo, de uma avaliação de imóvel, com área X, de pouco servirá estimar o inverso do valor do bem. O avaliador certamente reconstituirá o valor Y_c, procurado, promovendo nova inversão em $1/Y_c$. Neste momento, na realidade, estar-se-á estimando o valor do bem não mais pela equação I, mas por seu inverso, dado por

$Y_c = 1/(a + bX)$. (III)

Mas, se assim for, uma equação que não é mais linear passa a instrumentar a avaliação e surge, inevitavelmente, a pergunta: Nesta nova equação, se invocados os paradigmas de mínimos quadrados, os parâmetros a e b terão os mesmos valores?

Para respondê-la, basta retomar a busca de a e b, pelo mesmo critério de otimização adotado acima: escrever a equação da soma dos quadrados dos resíduos; obter as derivadas parciais da mesma, em relação aos parâmetros a e b, igualando-as, a seguir, a zero. (*) Assim, é reeditada a condição de "**mínimos quadrados",** a única forma de assegurar à equação resultante a confiabilidade que daí decorre. Este procedimento leva, novamente, a duas equações que permitem determinar a e b:

$$S(YY_c^2) = S(Y_c^3) \quad \text{(IV)}$$
$$S(XYY_c^2) = S(XY_c^3),$$

onde, evidentemente, Yc é dado pela equação III.

O recurso "SOLVER" das planilhas eletrônicas (e dos "Softwares" voltados para aplicações matemáticas) está habilitado a procurar os valores de a e b que satisfaçam estas condições. Alternativamente, para quem não se dispuser a procurar as derivadas parciais, há o recurso mais sumário de, simplesmente, submeter ao "SOLVER" a soma dos quadrados dos resíduos, $S(Y-Y_c)^2$, adotando, por exemplo, inicialmente, os valores de a e b fornecidos pela regressão linear e determinando a busca iterativa dos parâmetros capazes de minimizar esta soma, confiando que o resultado alcançado, ainda que não comprove a anulação das derivadas parciais, há de se situar muito próximo desta condição.

Se oura for a transformação de Y, análogo será o tratamento. Se, por hipótese, em lugar do inverso, Y for substituído por seu logaritmo natural, a equação I, acima, ficará

$$\ln Y_c = a + bX;$$

as equações II serão:

$$S(\ln Y) = Na + bSX$$
$$S(X\ln Y) = aSX + bSX^2;$$

a equação III terá a seguinte configuração "destransformada":

$$Yc = e^{(a = bX)};$$

as equações IV decorrerão das derivadas parciais de Z, soma dos quadrados dos resíduos, em relação a a e b, igualadas a zero, levando às seguintes condições que deverão satisfazer os parâmetros, para que o ajuste cumpra a exigência de mínimos quadrados:

$$SYY_c = SY_c^2$$
$$SXYY_c = SXY_c^2$$

Um exemplo

Tentando avançar no exame do assunto, os Autores, a partir de uma amostra de onze elementos (ver Tabela 1). com uma variável independente (X=Área) e uma variável dependente (Y=Valor unitário), examinaram os modelos de regressão que resultam das nove combinações possíveis entre três configurações das duas variáveis respectivamente inalterada, substituída por seu logaritmo neperiano ou por seu inverso. Somente três destes modelos apresentam coeficientes de determinação com valor superior a 0,70; neles, a variável X permanece inalterada e a variável dependente assume, respectivamente, a forma inalterada, a transformação pelo logaritmo neperiano e o valor inverso, alcançando os coeficientes de determinação de,

(*) Se necessário, recorrer a tabelas e formulários de matemática, ou a "Softwares" especializados em matemática.

Tabela 1

Dado	Área	V. unit.
1	5500	130
2	5000	150
3	4000	200
4	3500	250
5	2800	300
6	2000	400
7	1800	500
8	1600	550
9	1300	700
10	300	400
11	700	500

respectivamente, 0,707, 0,8398 e 0,8732 (ver Tabela 2). As seis combinações restantes foram descartadas, pois, além de alcançarem valores pouco expressivos nos coeficientes de determinação, as alterações da variável independente não trazem implicações para o objetivo do presente trabalho.

Tabela 2

Modelo	Forma X	Forma Y	Equação Linear	Coef. Det
1	X	Y	y = a + bx	0,707
2	X	Ln y	lny = a + bx	0,8398
3	X	1/y	1/Y=a+bX	0,8732

Normalmente, o que ocorreria? Provavelmente, o avaliador tomaria o modelo 3 como o mais confiável e o adotaria para avaliar o inverso do valor unitário do bem que constitui o objeto da avaliação. A seguir, desfazendo a inversão, teria o valor unitário procurado.

No entanto, assim fazendo, a equação adotada para avaliar o valor unitário seria, efetivamente, o inverso da equação linear processada na regressão, não mais a equação I, acima, mas a III. Todavia, embora operando na equação III, a solução anunciada utilizou os parâmetros a e b da equação I, arbitrariamente invocados, sem perguntar se, agora, com a inversão da equação a que se destinavam, ainda asseguram o cumprimento do preceito de "mínimos quadrados".

Evidentemente, não o fazem. Para tanto, precisariam satisfazer as condições expressas nas duas equações identificadas em IV. Os valores que as satisfazem são bem diversos, como a seguir se verá.

O mesmo ocorre, embora em grau menos pronunciado, com o Modelo 2, onde a variável dependente é substituída pelo seu logaritmo neperiano. O coeficiente de determinação supera, também, aquele da formulação inicial, não transformada. A Tabela 3, abaixo, contém os dados necessários à avaliação do modelo:

Tabela 3

Equação	a	b	R^2, Pseudo R^2
lnY=a+bX	6,553834877	-0,000295344	0,839889883
$Y=e^{(a+bX)}$	6,553834877	-0,000295344	0,595621088
$Y=e^{(a+bX)}$	6,44930244	-0,000227	0,637557241

A primeira linha contém os parâmetros a e b, bem como o coeficiente de determinação, que resultaram da regressão linear, com a variável Y transformada em lnY; a segunda linha transfere para a função não-linear, que exprime o valor de Y, os parâmetros lineares da solução linear e calcula o "Pseudo-R^2". Finalmente, na terceira linha, figuram os valores que a e b (obtidos com o emprego da

ferramenta “Solver”) deveriam assumir, para que a equação não-linear oferecesse uma solução de “mínimos quadrados”, com o respectivo valor de “Pseudo-R^2”. Com a solução “otimizada”, este coeficiente cresce em cerca de 7% em relação ao anterior, pois, como não poderia deixar de ser, melhora o ajuste da equação à tendência dos dados.
Na Tabela 4, abaixo, figura, para o Modelo 3, o mesmo tratamento oferecido pela Tabela 3 ao Modelo 2. Na primeira linha figuram os dados da regressão linear, com a transformação da variável Y, pelo inverso; na segunda linha, os valores obtidos para a função não-linear, resultante do desfazimento da inversão de Y, adotando, porém, os mesmos parâmetros já determinados na feição linear, anterior, tal como ocorre quando o avaliador opta pelo modelo transformado e, obtido 1/Y, calcula Y. A terceira linha contém os valores de a e b, que satisfazem as condições IV, obtidos através do emprego da ferramenta “Solver”, disponível em planilhas eletrônicas**.**

Tabela 4

Equação	**a**	**b**	**R^2, Pseudo R^2**
1/Y=a+bX	0,000595585	1,14E-06	0,87319
Y=1/(a+bX)	0,000595585	1,14E-06	-0,75056
Y=1/(a+bX)	0,001581678	4,96E-07	0,5439355

O valor do coeficiente de determinação, na quarta coluna, só é chamado de R^2 quando o modelo, além de linear, tem parâmetro linear diferente de zero. (1ª linha). Nos demais casos, é denominado de “Pseudo R^2” (Ref. 4), calculado pela expressão**:**

$$R^2 = 1 - SQR/SQT, \qquad (V)$$

em que SQR é a soma dos quadrados dos resíduos, na amostra, e SQT a soma dos quadrados das diferenças entre os valores de Y e a respectiva média. Na regressão linear, evidentemente, este resultado jamais será negativo, pois, necessariamente,

$$SQR < SQT,$$

já que:

$$SQT = SQR + SQE,$$

em que SQE é a soma dos quadrados dos afastamentos entre as ordenadas da equação de estimativa e a média dos Y.

Não obstante, na segunda linha da Tabela 4, chama atenção o fato de ser negativo o valor de “Pseuso-R^2”. Como pode tal ocorrer? A rigor, como já mencionado, o coeficiente de determinação nunca poderia ser negativo, pois, por definição, é o quadrado do coeficiente de correlação, cujo valor se situa entre –1 e + 1. O fato mostra quão ilusória é a superioridade do Modelo 3, onde o valor unitário foi substituído pelo seu inverso, levando ao coeficiente de determinação, na regressão linear, de 0,8732. Justificam-se, assim, plenamente, as recomendações de um especialista, quanto às cautelas a adotar no uso do coeficiente de determinação. (Ref. 4).

Desfeita a inversão, mas mantidos os parâmetros da regressão linear, ter-se-á uma equação não-linear, na qual pode ocorrer, nos termos da equação V, que a soma dos quadrados dos resíduos supere o valor da soma dos quadrados dos afastamentos dos valores de Y em relação à sua própria média. Em outras palavras: um ajuste tão defeituoso que seria preferível adotar, como estimador, simplesmente, a média dos valores da variável dependente!

Instado a comentar o fato pelos autores do presente trabalho, o Prof. Dr. Schabenberger (Ref. 5), afirmou:

"You are absolutely correct in that this measure can take on negative values in a non-linear model. This has never happened to me in practice and I have fit many, many nonlinear models. Hence, I do not consider that a shortcoming of the measure itself. Rather, a negative Pseudo-R^2 should indicate a serious problem with the model. I doubt, that if the appraiser would have fitted the model in its nonlinear form that it would have fitted the data at all. Sounds like poor model/data agreement to me. My experience is that if the nonlinear model is chosen properly, the Pseudo-R^2 will be sufficiently far away from 0 that negative values are not an issue.".

Finalmente, na terceira linha da Tabela 4 figuram os novos valores dos parâmetros da equação não-linear e o valor correto do coeficiente "Pseudo R^2", obtidos a partir da recondução do modelo à condição de mínimos quadrados, recorrendo à ferramenta "Solver" para buscar os valores de a e b que satisfaçam as equações IV, acima.

De qualquer forma, após esta análise, conclui-se que, apesar da aparência inicial em contrário, a solução linear, não transformada, $Y = a + bX$, com coeficiente de determinação de 0,707, prevalece sobre as duas transformações que pareciam superá-la.

Torna-se evidente, então, que a estimativa do valor do objeto da avaliação, quando, para fins de linearização, a variável dependente é transformada, corre o risco de ser obtida a partir de equação tecnicamente inaceitável, que não satisfaz o critério básico e fundamental de mínimos quadrados.

Valor unitário e valor total

Examinando o presente trabalho, poderia alguém perguntar: Se é tão pronunciado o zelo dos Autores em preservar o método de mínimos quadrados, por que, então, constróem seus exemplos com um modelo cuja variável dependente é o valor unitário, não o total? a passagem do primeiro para este último não implica, também, na modificação da equação estimadora, sem alteração dos seus parâmetros?".

À guisa de resposta, os Autores, além de admitir, desde logo, "**Boa pergunta**!", diriam que, a rigor, o uso do valor unitário não é visto como uma transformação de variáveis, mas como uma prática usual na definição da variável dependente. A amostra descrita na Tabela 1 foi colhida na prática profissional e assim foi aceita.

Todavia, o crítico está certo em apontar um fato inegável:. o emprego da variável dependente expressa pelo valor unitário, em lugar do valor total do bem avaliado, não deixa, também, de certa forma, de constituir-se num procedimento de transformação de variáveis. Neste caso, a estimativa se faz sobre o valor do bem, por unidade de área. A seguir, alcançada a estimativa, procede-se, como nos demais casos de transformação, ao desfazimento da alteração, para alcançar o valor total do objeto da avaliação. Ora, desfazer a alteração é multiplicar o valor unitário pela área, em geral a variável independente, no modelo adotado, modificando, em conseqüência, a função avaliadora, embora mantendo os parâmetros lineares antes determinados, sem perquirir se, agora, com a nova feição da equação, eles mantêm, ainda, o atributo de mínimos quadrados.

Seja um modelo expresso por

$$V_u=a+bX, \qquad \text{(VI)}$$

em que X é a área e V_u o valor unitário do bem avaliado. Obtido o resultado da regressão linear, o valor total do bem, V_t, é obtido por

$V_t = Vu * X$ (VII)

Invocando a equação VI, tem-se, então,

$V_t = aX + bX^2$ (VIII)

Neste caso, a avaliação final se faz por um trinômio incompleto do segundo grau; incompleto porque lhe falta o parâmetro linear; o parâmetro linear e o angular da regressão linear anterior passam a ser os coeficientes das variáveis X e X^2. Esta acolhida dos parâmetros não é precedida de qualquer verificação sobre se satisfazem, na nova equação (VIII), os requisitos da condição de mínimos quadrados. Esta é, no entanto, apenas a primeira crítica ao modelo; a segunda diz respeito à ausência do parâmetro linear. Com efeito, ainda que a e b assegurassem a condição de mínimos quadrados na equação VIII. Restaria, ainda, perguntar: por que impor que a otimização se proceda sobre uma equação forçada a passar pela origem dos eixos coordenados, sem a liberdade de procurar a solução ótima, deslocando-se, paralelamente, para interceptar o eixo das ordenadas, acima ou abaixo da origem, em busca de um mínimo absoluto na soma dos quadrados dos resíduos?

Poder-se-ia dizer: "A imposição é procedente, pois, se a área do objeto é nula, nulo também há de ser seu valor". Ocorre que o propósito da regressão por mínimos quadrados é o de determinar a equação que percorre o espaço ocupado pelos valores da amostra, de modo a minimizar a soma dos quadrados dos resíduos, **dentro deste intervalo**, isto é, para todos os valores de X compreendidos entre os limites da amostra, não se cogitando de extrapolações, como seria a avaliação da área nula, não contemplada na amostra.

Veja-se um exemplo:

Tabela 5

Dado	**Área**	**Valor Unit.**
1	1360	88
2	1740	201
3	2310	395
4	2595	354
5	2785	334
6	2975	316
7	3735	254
8	4400	218
9	4875	198
10	5825	166
11	6300	154

A regressão linear para estimar, diretamente, o valor total, é pouco expressiva, com $R^2=0,4211$; para o valor unitário, obtêm-se os resultados assinalados na primeira linha da Tabela 6, a seguir:

Tabela 6

Modelo	Equação	a	b	c	R^2
4	Vu=a+bX	308,2954	-0,01834	-	0,09811
5	$Vt=aX+bX^2$	308,2954	-0,01834	-	0,49536
6	$Vt=aX+bX^2$	405,8969	-0,04045	-	0,69141
7	$Vt=a+bX+bX^2$	-586404,9011	737,0418	-0,08060	0,80434

No Modelo 5 acha-se a avaliação do valor total, com os mesmos parâmetros do Modelo 4, mas multiplicando, em cada caso, o valor unitário, ali referido, pela respectiva área. No Modelo 6 figura o resultado da otimização do Modelo 5, isto é, com os valores dos parâmetros a e b decorrentes da imposição da soma mínima do quadrado dos resíduos. O valor de "pseudo-R^2" salta de 0,4975 para 0,6923. Finalmente, no Modelo 7, figura o trinômio completo do segundo grau, com reintrodução do parâmetro linear, elevando o coeficiente de determinação para 0,8048.

Conclusão

O presente trabalho está centrado numa preocupação dominante: O empenho em fazer com que a prática da avaliação, quando invocada a metodologia da regressão, mantenha-se estritamente aderente à condição de mínimos quadrados, sob pena de oferecer precedentes para resultados duvidosos e, inclusive, artifícios inaceitáveis de manipulação de resultados.

O problema reside na prática adotada, de procurar, pela transformação de variáveis, obter uma versão mais linearizada do modelo. Nada há a objetar quando esta transformação atinge apenas as variáveis independentes; mas, quando aplicada à variável dependente, tudo se passa como se instituída fosse, na busca da solução final, uma etapa preliminar, formada pela otimização da variável transformada. Cumprida esta etapa, desfaz-se, então, a transformação da variável, para obter o valor da grandeza que constitui o objetivo real da avaliação. Mas, ao assim fazer, são mantidos os valores dos parâmetros da equação linear da etapa preliminar, quando, na etapa final, a equação estimadora deixa de ser linear e a minimização da soma dos quadrados dos seus resíduos exigiria a busca de novos parâmetros otimizadores.

Em razão disto, os Autores propõem que, ao empregar modelos *linearizantes* para a regressão, o perito tenha a cautela de adotar, sempre, uma de duas regras: a) Nunca modificar a variável dependente, ou, b) Se o fizer, pesquisar novos valores para os parâmetros "a" e "b_i" (i = 1, 2, 3... n; "n" o número de variáveis independentes), de forma a obter um modelo não-linear aderente ao requisito de mínimos quadrados. Proposição similar fazem quanto ao emprego do valor unitário como variável dependente: Caso adotado este procedimento, recomenda-se, igualmente, ao calcular o valor total, buscar, novamente, na função resultante do produto do valor unitário pela área, os novos valores dos parâmetros otimizadores. Sugerem, ademais, que seja reintroduzindo o parâmetro linear na equação final, para melhorar seu poder estimador, pois, na multiplicação da equação do valor unitário pela área, este parâmetro passa a ser coeficiente da variável independente.

Sugerem os Autores, ademais, que, após exame mais detido da matéria, seja, eventualmente, examinada a hipótese de se incorporarem disposições disciplinadoras de tais procedimentos no texto da Norma aplicável a avaliações e perícias.

REFERÊNCIAS

(1) PIRES DA SILVA, Sérgio Alberto, “Seleção de Modelo por Variação Residual”. Anais do 2º Simpósio Brasileiro de Engenharia de Avaliações, “AVALIAR 2000”, ABDE, setembro de 2000, São Paulo.
(2) SPIEGEL, M. R. & LEMES DA SILVA, J. M., “Manual de Fórmulas, Métodos e Tabelas de Matemática, MAKRON *Books* do Brasil Editora, Ltda. São Paulo, 1992.
(3) WOLFRAM, STEPHEN, “*The Mathematica Book*”, 4th ed., Wolfram Media/Cambridge University Press, 1999.
(4) KVALSETH, Tarald O., “”Cautionary Note About R2 “. The American Statistician, November 1985, vol. 39, p. 279-285.
(5) SCHABENBERGER, Oliver, “Nonlinear Regression with the SAS system”, 1998. “Site” www.stat.vt.edu/~oliver, Item 3.5, “Calculating a R^2-type Measure”.
(6) SCHABENBERGER, Oliver, *e-mail*, 24/10/2000.