Nelineární regrese a tvorba matematických modelů

Jiří Militký, KTM, VŠST Liberec, 461 17 LIBEREC

1. Úvod

Nelineární regrese se v technických aplikacích používá velmi často při konstrukci matematických modelů. Pro tyto účely však již obecné nepostačuje klasické kriterium nejmenších čtverců odchylek (MNČ), ale je třeba modelovat i struktury chyb.

V tomto příspěvku je nejdříve stručně pojednáno o jednotlivých fázích matematického modelování s využitím regresních metod. Druhá část je zaměřena na využití některých technik transformace a vážení pro zlepšení rozdělení chyb, resp. snížení vlivu nelinearity modelu na statistickou analýzu odhadů.

2. Modelování s využitím regresních metod

Pod pojmem modelování se obecně chápe záměrná činnost přiřazující zkoumanému systému (originál objekt) jiný známý systém (fyzikální, abstraktní) zvaný model. Při konstrukci modelu se využívá vhodných kriterií vyjadřujících míru shody s jistým rysem originálu. V řadě praktických případů se konstruují modely matematické, které lze relativně snadno použít (např. pro optimalizaci atd).

Vlastní postup modelování je výrazně ovlivněn typem modelovaného systému.

Relativně nejjednodušší je konstrukce modelů pro nesložité, malé a dobře organizované systémy (laboratorní experimenty v idealizovaných podmínkách). Modely těchto systémů vycházejí z fyzikálně chemických zákonů a závisí na relativně malém počtu parametrů s definovaným fyzikálním významem. Je obyčejně také známo, v jakém rozmezí se mohou parametry pohybovat. Pokud lze závislost mezi výstupní (vysvětlovanou) a vstupními (vysvětlujícími) proměnnými znázornit, jsou patrné nejen trendy , ale také lokální tvarové zvláštnosti. V anglosasské literatuře se často tento

typ úloh označuje jako "curve fitting".

Je zajímavé, že také u velkých, špatně organizovaných systémů (typických pro behavioristické vědy) není postup konstrukce modelů relativně složitý. Vzhledem k existenci řady proměnných a nezachytitelných vlivů se často používá modelů ve tvaru Taylorova rozvoje neznámé funkce obyčejně do maximálně kvadratických členů. Pro zjištění významnosti jednotlivých vysvětlujících proměnných a jejich interakcí tento postup často vyhovuje. Jde o klasický případ regrese (v řadě úloh lineární).

Nejvíce komplikované z hlediska modelování jsou vědy technické (průmyslové experimenty). Zde se obyčejně pracuje s hůře organizovanými "difusními" systémy, kde sice platí fyzikálně chemické zákony, ale jejich vliv je "maskován" méně známými, nepopsanými vlivy. Datové orientované modely typické pro velké i špatně organizované systémy neumožňují stanovení modelových parametrů fyzikální povahy. Na druhé straně jsou však modely vycházející z fyzikálních představ silně zjednodušené a nepopisují obyčejně dobře chování originálu. Typickým jevem je zde multimodelovost a vlastní modelování se stává téměř "uměním".

V mnoha případech se postup modelování mění v závislosti na cíli modelování. Je zřejmé, že modely pro účely predikce mohou být tvořeny jako datově závislé (kriteriem je shoda modelu s experimentem). Na druhé straně modely strukturní musí být v souladu s hypotézami a teoriem o modelovaném systému.

Omezme se v dalším na jednu speciální třídu modelů, kterou lze označit jako deterministické modely s náhodným rušením.

Vysvětlující proměnné $\underline{x}$, nechť jsou deterministické. Tyto proměnné při působení na orginál vyvolají měřitelný důsledek y*, který lze vyjádřit deterministickým modelem

$$y^* = f(\underline{x}, \underline{a}) \qquad (1)$$

kde f(.) je modelová funkce a $\underline{a}$ jsou modelové parametry. Důsledkem způsobu měření veličiny y*, kolísání nastavených podmínek a nezachytitelných vlivů je výsledná vysvětlovaná

veličina y náhodná

$$y = G(y^*, e) \qquad (2)$$

Zde symbol e označuje souhrnnou chybu (považuje se za náhodnou veličinu charakterizovanou např. hustotou pravděpodobnosti p(e)) a G(.) vyjadřuje pravděpodobnostní model působení poruch.

Obecně je tedy třeba při modelování těchto typů úloh hledat deterministický model $f(\underline{x},\underline{a})$ a pravděpodobnostní model p(e). Ve velké většině případů se v praxi věnuje pozornost pouze specifikaci modelu $f(\underline{x},\underline{a})$ s tím, že se předpokládá aditivní model působení poruch

$$G(f(\underline{x},\underline{a}), e) = f(\underline{x},\underline{a}) + G \cdot e_s \qquad (3)$$

Zde je směrodatná odchylka, e_s je náhodná veličina s nulovou střední hodnotou a jednotkovým rozptylem. Pokud je $p(e_s)$ hustota pravděpodobnosti normovaného normálního rozdělení lze aplikací metody maximální věrohodnosti dospět ke kritériu MNČ. Neplatí-li však model (3), může být použití MNČ zdrojem toho, že není jinak korektní model $f(\underline{x},\underline{a})$ prakticky akceptován.

Z toho plyne, že podcenění konstrukce pravděpodobnostního modelu je jednou z příčin možného nezdaru celého procesu modelování.

3. Postup při modelování

Při modelování deterministických systémů s náhodným rušením se běžně vychází z experimentu provedeného na originálu. Výsledkem je N-tice hodnot vysvětlované (výstupní) proměnné y_i, pro zvolené nastavení vysvětlujících proměnných $\underline{x}_i$. Tyto výsledky experimentu lze chápat jako N-tici bodů v E^{m+1} rozměrném prostoru.

Vlastní postup modelování se obecně skládá z těchto fází:

1. Definice úlohy
2. Realizace experimentu
3. Konstrukce deterministického modelu
4. Konstrukce pravděpodobnostního modelu

5. Adjustace modelu

6. Verifikace modelu

V každé fázi záleží zejména na znalostech o modelovaném systému a zkušenostech modelujícího. Platí zde v plné míře výrok "Experimentátor ví vždy o problému více než jeho počítač". Na druhé straně právě použití počítače umožňuje zahrnutí obecnějších typů pravděpodobnostních modelů a diskriminací mezi nimi.

3.1. Definice úlohy

Tato etapa se považuje velmi často za zcela triviální. V mnoha případech je však právě nejasná definice cíle modelování zdrojem řady problémů. Např. úloha "vyhlazení modelu" za účelem následné numerické derivace, resp. integrace modelu (častá třeba v chemii) se dá jednoduše řešit s využitím regresních spline modelů, vyjádřených jednoduše ve tvaru useknutých polynomů [4]

$$E(y/x) = \sum_{j=0}^{m} a_j x^j + \sum_{L=1}^{k} b_L (x - t_L)_+^m \qquad (4)$$

Zde t_L jsou tzv. uzlové body (místa styku lokálních polynomů stupně m). Symbol $(x)_+$ je useknutá funkce, pro kterou platí

$$(x)_+ = x \qquad \text{pro } x > 0$$
$$(x)_+ = 0 \qquad \text{pro } x \leq 0$$

Je zřejmé, že pro zvolené t_L L = 1.. k je úloha odhadů parametrů (a_j, b_L) j = 1 .. m, L = 1 .. k úlohou lineární regrese. Při interaktivním způsobu volby uzlových bodů na displeji lze snadno modelovat i velmi kmplikované polymodální závislosti, pro které je nalezení matematických modelů jinak takřka nemožné.

Také v situacích, kdy je třeba hledat optimální podmínky funkce systémů (úloha extremalizace) postačuje místo konstrukce složitých modelů použití polymonů nízkého stupně, případně spline polynomů.

V řadě praktických úloh tedy může správná specifikace cíle značně ulehčit vlastní modelování.

3.2. Realizace experimentu

Obecné je třeba při vlastní realizaci experimentu stanovit experimentální oblast (obyčejně jako rozmezí všech vysvětlujících proměnných) a způsob jejího "zmapování", tj. určení poloh jednotlivých složek $\underline{x}_i$.

Platí pravidlo, že dobře naplánovaný experiment značně ulehčí jak numerickou, tak i statistickou stránku vlastní regrese (odhadu parametrů). Na druhé straně řada potíží, zejména při více vysvětlujících proměnných tvoří multikolinearita jako typický případ špatně plánovaného experimentu.

Z celé řady kritérií pro vhodné plánování experimentů, uveďme pouze tzv. D-optimalitu, kdy se minimalizuje determinant matice $(J^TJ)^{-1}$. Matice J má prvky

$$J_{ij} = \frac{\partial f(\underline{x}_i, \underline{a})}{\partial a_j} \qquad \begin{matrix} i = 1 \dots m \\ j = 1 \dots m \end{matrix} \tag{5}$$

Obyčejně platí, že pro špatně postavené modely $f(\underline{x},\underline{a})$ nevede konstrukce D-optimálního experimentálního plánu k výraznému zlepšení. V řadě praktických úloh se také zpracovávají výsledky již hotových (neoptimálně plánovaných) experimentů.

V těchto situacích nezbývá než vzít v úvahu možné důsledky způsobené např. multikolinearitou [4].

3.3. Konstrukce deterministického modelu

I když existují jistá pravidla pro obecné řešení úlohy modelování (uvedená např. v Descartově spisu "Pravidla o řízení rozumu") nelze nalézt universální postup.

Pro případ konstrukce třídy empirických modelů založených na geometrické podobnosti modelu s experimentem existuje celé spektrum technik využívajících různých katalogů nebo atlasů křivek a ploch. S oblibou se používají takové funkce, které lze linearizovat nebo jejichž tvar závisí jednoduše na hodnotách parametrů. Pro konstrukci vícerozměrných empirických modelů lze použít také speciální metody regresní diagnostiky (grafy konstrukční proměnné, přidané proměnné atd.) [4].

Obecně platí, že v praxi se této fázi modelování věnuje maximální pozornost, protože jde o "viditelný" výsledek celé činnosti.

3.4. Konstrukce pravděpodobnostního modelu

Tato etapa modelování je v praxi běžně zanedbávána. Navíc se konstrukce pravděpodobnostního modelu projeví "jen" na odhadech parametrů a jejich statistickém chování.

Jde tedy o "skrytou" část modelu, která však může celý proces modelování rozhodujím způsobem ovlivnit. Při konstrukci pravděpodobnostního modelu G(.) lze postupovat třemi způsoby.

A. Sestavit teoretický model působení chyb na základě znalostí a hypotéz o jejich vzniku a statistickém chování.

B. Sestavit empirický, velmi obecný model zahrnující většinu modelů působení chyb a hledat jeho zjednodušení pro daný případ.

C. Na základě ověření předpokladů MNČ (po její aplikaci) provést jejich zpřesnění. To se týká zejména heteroskedasticity a autokorelačních struktur.

Ve druhé části tohoto příspěvku jsou detailněji rozebrány cesty A a B. Postup ad C je vlastně iterativní aplikace metod regresní diagnostiky, která je popsána např. v práci [4].

3.5. Adjustace modelu

V této etapě se provádí odhad modelových parametrů a jejich statistická analýza (testy hypotéz).

Oblíbený je Bayesovský přístup, který lze často v kontextu modelování chápat jako zpřesňování à priorních informací s využitím informací získaných z experimentu.

Pomocí Bayesovy definice podmíněné pravděpodobnosti lze à posteriorní hustotu pravděpodobnosti $p(\underline{a}/y)$ vyjádřit pomocí à priorní hustoty pravděpodobnosti modelových parametrů $p(\underline{a})$ a věrohodnostní funkce $L(\underline{a})$ ve tvaru

$$p(\underline{a}/y) = \frac{L(\underline{a}) \cdot p(\underline{a})}{p(y)} \qquad (6)$$

Bayesovský odhad $\underline{a}^*$ pak maximalizuje a posteriorní hustotu pravděpodobnosti, tedy

$$\underline{a}^* = \max \; p(\underline{a}/y) \tag{7}$$

Pokud nejsou k dispozici a priorní informace o parametrech je Bayesovský odhad $\underline{a}^*$ shodný s odhadem maximalizujícím věrohodnostní funkci

$$\underline{a}^* = \max L(\underline{a}) \equiv \max \ln L(\underline{a}) \tag{8}$$

Za předpokladu, že y_i jsou nezávislé, stejně rozdělené náhodné veličiny s hustotou pravděpodobnosti $p(y_i/\underline{a})$ má věrohodnostní funkce tvar

$$L(\underline{a}) = \prod_{i=1}^{N} p(y_i/\underline{a}) \tag{9}$$

Pokud platí model měření vyjádřený rov. (2) a chyby e mají hustotu pravděpodobnosti $p_e(y_i/\underline{a})$, lze snadno určit, že

$$p(y_i/\underline{a}) = p_e\left[G^{-1}(f(\underline{x}_i, \underline{a}), y_i)\right] \left| \frac{\partial G^{-1}(.)}{\partial y_i} \right| \tag{10}$$

kde $G^{-1}(.)$ je formálně inverzní funkce k funkci $G(.)$. Je zřejmé, že pro aditivní model měření

$$G(.) = y^* + e$$

je

$$G^{-1}(.) = y - y^*$$

a tedy

$$p(y/\underline{a}) = p_e[y - f(\underline{x}, \underline{a})] \tag{11}$$

Pro multiplikativní model měření

$$G(.) = y^* \cdot \exp(e)$$

je inverse

$$G^{-1}(.) = \ln y - \ln y^*$$

Po dosazení do rov. (10) pak vyjde

$$p(y_i/\underline{a}) = \frac{1}{|y_i|} \cdot p_e\left[\ln y_i - \ln f(\underline{x}_i, \underline{a})\right] \qquad (12)$$

Pro případ, že $p_e(.)$ je hustota pravděpodobnosti normálního rozdělení $N(0,\sigma^2)$, vyjde po dosazení z rovnice (11) do rovnice (9) a (8), že maximálné věrohodné odhady minimalizují kriterium MNČ

$$S(\underline{a}) = \sum_{i=1}^{N} \left[y_i - f(\underline{x}_i, \underline{a})\right]^2 \qquad (13)$$

Podobné resultuje pro multiplikativní model chyb (rov. (12)) kriterium MNČ v logaritmech.

Pokud nejsou chyby e nezávislé, ale jsou charakterizovány kovarianční maticí C_e předpokládá se obyčejné, že mají vícerozmérné normální rozdělení. Modeluje se pak pouze kovariační matice.

Z výše uvedeného je patrné, že v řadě případů vede adjustace modelu na problém minimalizace součtu čtverců. Pro tuto úlohu existuje řada speciálních algoritmů (Gauss-Newtonovy a Marquardtovy metody jsou nejznáméjší), které umožňují vlastní numerickou extremalizaci. Ve složitějších případech je třeba použít obecných optimalizačních metod vhodných pro maximalizaci sestavené věrohodnostní funkce. (Známý je např. program MAXLIK v jazyce GAUSS).

Při statistické analýze se většinou vychází z představy, že lze v okolí odhadu $\underline{a}^*$ funkci $f(\underline{x},\underline{a})$ dostatečně přesné linearizovat. Pak lze vyjádřit kovariační matici odhadů ve tvaru

$$D(a) = \sigma^2 (J^T J)^{-1} \qquad (14)$$

a provádět statistickou analýzu podle stejných vztahů jako u lineární regrese (při náhradé matice X maticí J).
Kvalita linearizace souvisí úzce s nelinearitou regresního modelu. Dá se také vyjádřit pomocí vychýlení odhadů $\underline{b}$. Z hlediska jednoduchosti vyjádření je výhodné počítat vychýlení dle Cooka a Tsaie

$$\underline{b} = (J^T J)^{-1} J^T \underline{d} \qquad (15)$$

kde $\underline{d}$ je vektor se složkami

$$d_i = -\sigma^2 \, tr\left[(J^T J)^{-1} \cdot G_i \right] / 2 \qquad (16)$$

V rov. (16) je tr(.) stopa matice a matice G_i má složky

$$G_{j,L} = \frac{\partial f(\underline{x}_i, \underline{\alpha})}{\partial \alpha_j \, \partial \alpha_L} \qquad j, L = 1 \ldots m \qquad (17)$$

Vychýlení odhadů se považuje za akceptovatelné pokud

$$| 100 \, b_j / \alpha_j^* | \leq 1$$

3.6. Verifikace modelu

V této etapě se ověřuje, zda model:

a) vyhovuje datům

b) vyhovuje à priorním představám o parametrech

c) má dostatečné predikční schopnosti.

V řadě případů se provádí také diskriminace mezi modely, při které se ještě zohledňuje to, aby byl model co nejjednodušší (s nejmenším počtem parametrů).

Existuje celá řada technik pro vyjádření shody modelu s daty a pro posouzení jeho predikčních schopností.
Jednoduše lze počítat střední kvadratickou chybu predikce dle vztahu

$$MEP = \frac{1}{N} \sum_{i=1}^{N} \left[y_i - f(\underline{x}_i, \underline{\alpha}^*_{(-i)}) \right]^2$$

kde $\underline{\alpha}^*_{(-i)}$ jsou odhady získané při vynechání i-tého bodu $(\underline{x}_i, y_i)$.

Zobecněním tohoto postupu jsou metody typu " Cross validation ", kdy se počítají odhady z "trénovacích" dat a chyba predikce z "ověřovacích" dat.

4. Pravděpodobnostní modely měření

Jak již bylo uvedeno lze při konstrukci modelu působení chyb využít buď ž priorních představ nebo sestavit obecný empirický model. Zde si ukažme obě tyto techniky.

4.1. Teoretický model chyb

Teoretické modely chyb se tvoří na základě konkretních znalostí o modelovaném systému a měřícím procesu. Většinou se tyto modely týkají typu heteroskedasticity nebo autokorelační struktury chyb. Je např. známo, že u řady přístrojů, které měří v rozsahu několika řádů nelze dodržet požadavek konstantního rozptylu měření σ^2. Je však často docíleno konstantnosti relativní chyby měření definované výrazem

$$s_R = \sigma / f(\underline{x}, \underline{a})$$

Rozptyl měření je pak modelován jako kvadratická funkce hodnoty regresního modelu

$$D(y_i) = \sigma_0^2 f^2(\underline{x}_i, \underline{a}) \qquad (18)$$

Pokud platí ostatní standardní předpoklady o chybách měření vede použití rovnice (18) k modelu

$$y = f(\underline{x}, \underline{a}) + \sigma_0 \cdot f(\underline{x}, \underline{a}) \cdot e_s \qquad (19)$$

kde e_s je opět standardizovaná náhodná veličina s nulovou střední hodnotou a jednotkovým rozptylem. Rovnice (19) je případem nelineárního heteroskedastického modelu, pro který vede použití metody maximální věrohodnosti k úloze minimalizace kriteria

$$S_W = \sum_{i=1}^{N} [y_i - f(\underline{x}_i, \underline{a})]^2 / f^2(\underline{x}_i, \underline{a}) \qquad (20)$$

Pokud se dá předpokládat, že chyby měření jsou malé, lze místo kriteria S_W použít jednodušší kriterium čtverců relativních odchylek

$$S_R = \sum_{i=1}^{N} \left[1 - f(\underline{x}_i, \underline{a}) / y_i \right]^2 \qquad (22)$$

Typickým případem vzniku speciální autokorelační struktury chyb je experimentování na jednom vzorku. Příkladem je např. sledování průběhu chemických reakcí, kdy se z reaktoru odebírají v určitých intervalech vzorky, které se analyzují na měřicím přístroji [4].

Celková chyba e_i se v těchto situacích skládá z chyb způsobených fluktuací podmínek procesu u_j (ve všech časech až do i-tého) a chyby měření v_i. To lze vyjádřit vztahem

$$e_i = \sum_{j=1}^{i} u_j + v_i \qquad (23)$$

Rov.(23) vyjadřuje fakt, že se chyby způsobené neidealitami procesu kumulují.

Pro jednoduchost předpokládejme, že

a) chyby procesu a chyby měření jsou vzájemně nezávislé
$E(u_j, v_i) = \emptyset \qquad j = 1 \ldots i$

b) chyby procesu jsou vzájemně nezávislé $E(u_i, u_j) = \emptyset$
Mají konstantní rozptyl $D(u_i) = \sigma_u^2$ a nulovou střední hodnotu $E(u_i) = \emptyset$

c) chyby měření v_i jsou také vzájemně nezávislé $E(v_i, v_j) = \emptyset$
Mají konstantní rozptyl $D(v_i) = \sigma_v^2$ a nulovou střední hodnotu $E(v_i) = \emptyset$.

Rov.(23) se dá vyjádřit také ve tvaru

$$e_i = e_{i-1} + w_i \qquad (24)$$

kde e_{i-1} je chyba v (i-1)ním čase a w_i je souhrnná chyba odpovídající přechodu systému z (i - 1)ního do i-tého stavu. Na základě výše uvedeného platí, že

$$E(w_i) = \emptyset \qquad D(w_i) = \sigma_u^2 + \sigma_v^2 = \tau^2$$

Chyby e_i však již nejsou nezávislé, a platí pro ně

$$e_i = \sum_{j=1}^{i} w_j \qquad e_1 = w_1$$

Maticově lze pak vektor chyb $\underline{e}$ vyjádřit ve tvaru

$$\underline{e} = A\underline{w} \tag{25}$$

kde A je dolní trojúhelníková matice s jedničkami na a pod hlavní diagonálou.
Pro kovarianční matici chyb $\underline{e}$ pak platí

$$C_e = E(\underline{e}\cdot\underline{e}^T) = \tau^2 A\cdot A^T \tag{26}$$

a její inverze je rovna

$$C_e^{-1} = \tau^{-2}(A^{-1})^T(A^{-1}) \tag{27}$$

Vzhledem ke speciální struktuře matice A je matice A^{-1} bidiagonální.
Diagonální prvky jsou rovny 1 a prvky poddiagonálního pásu jsou rovny -1. Ostatní prvky jsou nulové.
Pokud mají chyby $\underline{e}$ normální rozdělení je sdružená hustota pravděpodobnosti p(y/$\underline{a}$) rovna

$$p(y/\underline{a}) = (2\pi\tau^2)^{-\frac{N}{2}}\cdot\exp\left[-0{,}5(\underline{y}-\underline{f})^T C_e^{-1}(\underline{y}-\underline{f})\right] \tag{28}$$

kde vektor $\underline{f}$ má prvky $f(\underline{x}_i,\underline{a})$.
Vzhledem k tomu, že rovnice (28) definuje přímo věrohodnostní funkci, je zřejmé, že maximálně věrohodné odhady lze získat minimalizací výrazu v exponenciále. Ten je možno vyjádřit ve tvaru

$$S_L = (\underline{y}-\underline{f})^T C_e^{-1}(\underline{y}-\underline{f}) = \sum_{i=1}^{N}\left[L_i - K_i(\underline{a})\right]^2 \tag{29}$$

Dále platí

$$L_i = y_i - y_{i-1} \qquad K_i(\underline{a}) = f(\underline{x}_i,\underline{a}) - f(\underline{x}_{i-1},\underline{a})$$

$$y_0 = f(\underline{x}_0,\underline{a}) = \emptyset$$

Kriterium S_L tedy odpovídá metodě nejmenších čtverců pro <u>první diference</u>.

Opuštěním předpokladů normality, resp. aditivity chyb lze

dospět ke komplikovanějším modelům G(.), kterým odpovídají složitější kriteria regrese.
Tak pro případ modelování neizotermních kinetických procesů se ukázalo jako vhodné kriterium čtverců logaritmů prvních diferencí (kumulativní procesní chyby a multiplikativní model měření).
Složitější modely chyb zahrnující i chyby ve vysvětlujících proměnných jsou uvedeny v [1].

4.2. Empirický model chyb

Při konstrukci empirického modelu chyb je snahou nalézt takový výraz, který by zahrnoval pokud možno všechny typické struktury chyb.

S výhodou se zde využívá vhodné mocninné transformace, která může často zajistit zkonstantnění rozptylu a zlepšení symetrie rozdělení chyb (přiblížení k normalitě). Klasickým příkladem je Box-Coxova rodina transformací závisející na jednom parametru c. Tato rodina transformací je definována vztahy

$$z^{(c)} = (z^c - 1)/c \qquad c \neq 0$$

$$z^{(c)} = \ln z \qquad c = 0$$

S využitím této transformace lze vyjádřit celou třídu modelů měření ve tvaru

$$y^{(c)} = f^{(c)}(\underline{x}, \underline{a}) + G \cdot e_s \qquad (30)$$

Je zřejmé, že pro c = 1 rezultuje z rovnice (30) aditivní model měření a pro c = 0 multiplikativní model měření.
Model (30) má tu zajímavou vlastnost, že $f(\underline{x},\underline{a})$ představuje podmíněný medián veličiny y nezávisle na velikosti c [3]. Toto platí pro případ, že e_s v modelu (30) má symetrické rozdělení.

V práci [3] bylo ukázáno pomocí Taylorova rozvoje modelu (30), že tento model je možno vyjádřit také ve tvaru

$$y \approx f(\underline{x}, \underline{a}) + G f^{1-c}(\underline{x}, \underline{a}) \cdot e_s + O(G^2) \qquad (31)$$

Pro malé G je tedy model (30) ekvivalentní nelineárnímu heteroskedastickému regresnímu modelu.

Platí tedy např., že multiplikativní model měření (c = 0) je přibližné ekvivaletní modelu heteroskedasticity vyjádřenému rov (18) tj. případu konstantních relativních chyb měření.

V některých případech model (30) neumožňuje simultální přiblížení k normalitě (resp. symetrickému rozdělení) a odstranění heteroskedasticity.
Pak lze provést další rozšíření o funkci heteroskedasticity a dospět k modelu

$$y^{(c)} = f^{(c)}(\underline{x}, \underline{a}) + \sigma \cdot g(x,d)\, e_s \qquad (32)$$

Funkce g(x,d) se často volí ve tvaru [2]

$$g(x,d) = x^d$$

nebo ve tvaru [5]

$$g(x,d) = f^{d}(\underline{x}, \underline{a})$$

Komplikovanější typy g(x,d) jsou vhodné pro případ, kdy se uvažují také chyby ve vysvětlujících proměnných [1].
Lze dokázat, že model (32) odpovídá přibližně nelineárnímu heteroskedastickému regresnímu modelu

$$E(y) = f(\underline{x}, \underline{a})$$

$$D(y) = \sigma^2 g^2(x,d) \cdot f^{2(1-c)}(\underline{x}, \underline{a})$$

Za předpokladu, že v modelu (32) mají již chyby e_s přibližné normované normální rozdělení je možné vyjádřit logaritmus věrohodnostní funkce ve tvaru

$$\ln L = \sum_{i=1}^{N} \left\{ (c-1)\cdot \ln(y_i) - \ln[\sigma\, g(x_i,d)] \right\} - \frac{1}{2\sigma^2} \sum_{i=1}^{N} \left\{ [y_i^{(c)} - f^{(c)}(\underline{x}_i, \underline{a})] / g(x_i,d) \right\}^2 \qquad (33)$$

Pro pevné hodnoty parametrů c, d, a lze z rovnice (33) nalézt maximálně věrohodný odhad σ^2 analyticky

$$\hat{\sigma}^2(c,d,\underline{a}) = N^{-1} \sum_{i=1}^{N} \left\{ [y_i^{(c)} - f^{(c)}(\underline{x}_i, \underline{a})] / g(x_i,d) \right\}^2$$

Po dosazení do rov. (33) resultuje koncentrovaná věrohodnostní funkce

$$\ln L^* = \sum_{i=1}^{N} \left\{ (c-1) \ln y_i - \ln \left[\hat{G}(c,d,\underline{a}) \cdot g(x_i,d) \right] \right\} - N/2 \qquad (34)$$

Maximalizace rovnice (34) se dá provádět pomocí obecných extremalizačních programů nebo kombinací metody vážených nejmenších čtverců se systematickými změnami parametrů c, d. Po určení odhadů c^*, d^*, $\underline{a}^*$ je vhodné ověřit, zda nelze model vyjádřený rov.(32) zjednodušit.

Pro tyto účely je možné např. využít 100 $(1-\alpha)$ % ní oblasti spolehlivosti pro parametry (c, d) [5].

Označíme-li (c^*, d^*) odhady odpovídající maximu věrohodnostní funkce (34) můžeme zkonstruovat oblast spolehlivosti pro (c, d) hledáním všech dvojic (c_o, d_o), pro které platí, že

$$2 \left[\ln L^*(\hat{c},\hat{d}) - \ln L^*(c_o,d_o) \right] \leq \chi^2(2)$$

kde $\chi^2(2)$ je 100 $(1-\alpha)$ %ní kvantil chí kvadrát rozdělení.

V práci [2] je doporučeno použít pro testaci významnosti parametrů c, d speciálního F-testu.

Pro případ d = 0 existuje možnost jednoduššího přibližného odhadu parametru c (z linearizace obou stran modelu (30)) a testace jeho významnosti. Postačují přitom pouze výsledky nelineární regrese klasickou MNČ.

Transformace typu (30) obyčejně vede také k snížení vnitřní nelinearity regresního modelu.

Pro odstranění (nebo snížení) nelinearity způsobené parametry je buď možné provést reparametrizaci modelu nebo mocninnou transformaci odhadů $\underline{a}^*$. V obou případech lze s výhodou kontrolovat snížení nelinearity modelu pomocí vychýlení $\underline{b}$ definovaného rovnicí (15).