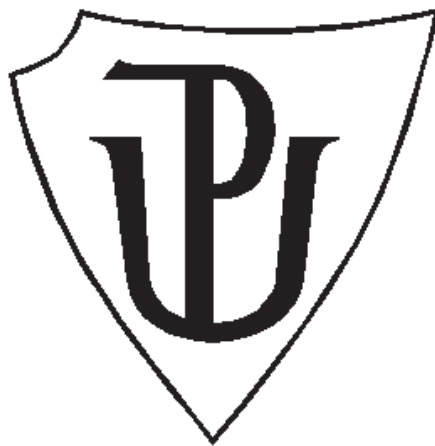


UNIVERZITA PALACKÉHO V OLOMOUCI
PŘÍRODOVĚDECKÁ FAKULTA
KATEDRA MATEMATICKÉ ANALÝZY A APLIKACÍ MATEMATIKY

SOUTĚŽNÍ PRÁCE

Konstrukce D-optimálního plánu měření
Langevinovy funkce a nelineární regresní modely

v oboru Aplikovaná matematika



Michaela Skopalíková

Školitelé: prof. RNDr. Ing. Lubomír Kubáček, DrSc., Dr.h.c.

Mgr. Jaroslav Marek, Ph.D.

Olomouc 2009

Obsah

1	Úvod	2
2	Formulace problému	4
2.1	Užití Levenberg–Marquardtova algoritmu	5
2.2	Design experimentu	6
3	Linearizované regresní modely	11
3.1	Algoritmus 1 – Lineární regresní model nepřímého měření vektorového parametru bez podmínek	11
3.2	Algoritmus 2 – Lineární regresní model přímého měření vektorového parametru se systémem podmínek typu II	13
3.3	Porovnání vhodnosti volby modelů	17
4	Závěr	18

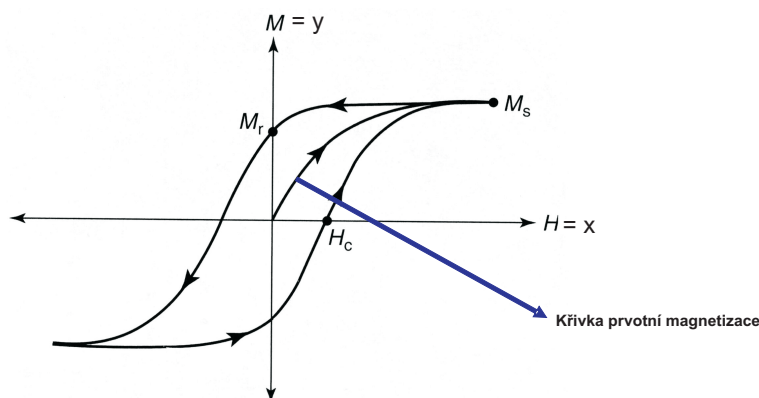
1 Úvod

Magnetické materiály díky velkému aplikačnímu potenciálu vzbuzují značný zájem odborné veřejnosti. Ukazuje se totiž, že bude-li se zmenšovat velikost magnetického materiálu, začne pod svou určitou, pro daný materiál charakteristickou, velikostí vykazovat zcela odlišné magnetické vlastnosti než jeho makroskopický protějšek. To je způsobeno tím, že magnetické chování v „nanosvětě“ je již řízeno jinými fyzikálními zákony, než v „makrosvětě“. V drtivé většině případů je pak možné pozorovat takové magnetické chování, které je velmi příznivé a přitažlivé pro jeho následnou aplikaci v praxi.

Z aplikačního hlediska jsou jedním z nejvýznamnějších druhů magnetických materiálů sloučeniny na bázi železa a oxidů železa. Jsou-li tyto sloučeniny navíc připraveny v nanometrovém rozměru, mají obrovské využití (např. pigmenty v známých médiích, kontrastní látky v nukleární magnetické rezonanci, oprava DNA kódů, cílená léčba rakoviny atd.).

Rozlišují se dva přístupy jak je možné syntetizovat nanočástice, fyzikální a chemický. Výsledkem obou těchto přístupů jsou různé systémy magnetických nanočástic, lišící se především svým rozměrem.

Jakmile je takovýto systém nanočástic připraven, je nutné zjistit jeho magnetické



Obrázek 1: Hysterézní smyčka a její významné body charakterizující zkoumaný magnetický materiál

vlastnosti, na základě kterých se poté rozhoduje, zda-li tento nasyntetizovaný systém splňuje požadavky kladené danou aplikací. Základní charakteristikou každého magnetického materiálu či nanomateriálu je magnetizace. Je to fyzikální veličina, která jednoznačně popisuje magnetický stav zkoumaného vzorku. Magnetizace materiálu nebo nanomateriálu se většinou měří buď při různé teplotě, ale s konstantním vnějším magnetickým polem, nebo při konstantní teplotě, ale s měnícím se vnějším magnetickým polem. Výstupem druhého zmíněného postupu měření je tzv. polní závislost magnetizace, která bývá obvykle označována jako hyste-

rezní smyčka. Z hysterezní smyčky je možné stanovit několik parametrů, které jednoznačně charakterizují daný materiál. Na základě hodnot těchto parametrů lze poté jednoznačně rozhodnout, zda-li je materiál vyhovující pro požadavky dané aplikace či nikoliv.

Jak již bylo řečeno, pro získání hysterezní smyčky je nutné zachovat při jejím měření konstantní teplotu a měnit sílu vnějšího magnetického pole.

Z matematického hlediska to znamená, že síla vnějšího magnetického pole je nezávisle proměnnou x , zatímco měřená magnetizace je závisle proměnnou y .

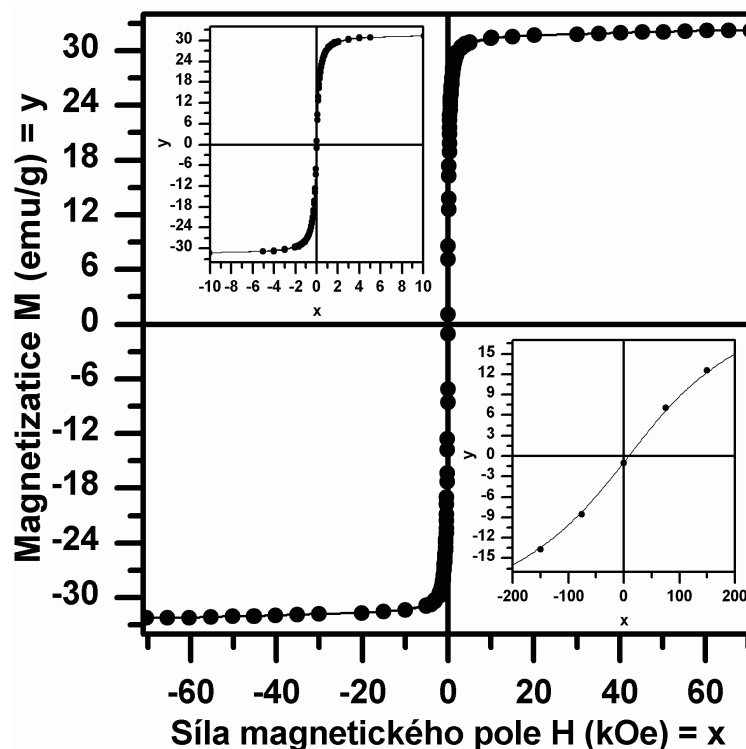
Na obrázku 1. je názorně zobrazen průběh hysterezní smyčky.

Nanometrové magnetické materiály vykazují jeden ze zvláštních fenoménů, který je právě stěžejní pro jejich praktické využití. Tento jev se nazývá superparamagnetismus a jeho hysterezní smyčka prochází počátkem (obrázek 2.).

Již v roce 1905 francouzský fyzik Pierre Langevin odvodil pro hysterezní smyčku v superparamagnetickém stavu závislost danou tímto předpisem

$$y = l_1 \cdot \coth(l_2 \cdot x) - \frac{l_1}{l_2 \cdot x},$$

kde parametry l_1 a l_2 jsou fyzikální konstanty charakterizující jednoznačně zkoumaný nanomateriál.



Obrázek 2: Hysterezní smyčka superparamagnetického nanometrového materiálu.

Naším cílem bude najít odhady parametrů této funkce pro data z fyzikálního experimentu. Odhad neznámých parametrů bychom mohli najít pomocí některých numerických algoritmů pro nelineární regresi. Tímto přístupem se však v této práci zabývat nebudeme. K aproximaci použijeme regresní model, když nejprve nelineární funkci pomocí Taylorova rozvoje převedeme na lineární aproximaci. V situaci, kdy experimentátor ještě neměřil, je náš cíl ještě větší. Máme navrhnout body, ve kterých má být měření uskutečněno tak, aby přesnost odhadu neznámých parametrů Langevinovy funkce byla co největší. To je náš nejdůležitější úkol.

V numerické části budeme aplikovat navržené algoritmy na data z měření nanočástic Gamma formy oxidu železitého, jejichž střední rozměr nabývá hodnoty přibližně 15 nm. Tyto nanočástice byly připraveny teplotní dekompozicí oktanu železnatého při teplotě 360° C. Experimentální měření bylo provedeno ve 150 bodech vybraných z intervalu -70000 Oe až 70000 Oe a bylo provedeno v Nanocentru Univerzity Palackého v Olomouci.

2 Formulace problému

Mějme naměřené hodnoty $(Y_1, Y_2, \dots, Y_N)'$ parametru β v bodech $(x_1, x_2, \dots, x_N)'$ určených deterministicky. Hodnoty x_i , $i = 1, \dots, N$, reprezentují stanovenou sílu vnějšího magnetického pole a hodnoty Y_i , $i = 1, \dots, N$, představují měření odpovídající magnetizace materiálu.

Hodnoty závislé proměnné mohou být vyjádřeny pomocí Langevinovy funkce takto:

$$Y_i = l_1 \cdot \coth(l_2 \cdot x_i) - \frac{l_1}{l_2 \cdot x_i}, \quad i = 1, \dots, N.$$

Model měření lze popsat

$$\mathbf{Y}_N \sim N_N \left[\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{pmatrix}, \Sigma \right] = N_N \left[\begin{pmatrix} l_1 \cdot \coth(l_2 \cdot x_1) - \frac{l_1}{l_2 \cdot x_1} \\ l_1 \cdot \coth(l_2 \cdot x_2) - \frac{l_1}{l_2 \cdot x_2} \\ \vdots \\ l_1 \cdot \coth(l_2 \cdot x_N) - \frac{l_1}{l_2 \cdot x_N} \end{pmatrix}, \Sigma \right].$$

Z neznámých parametrů vytvoříme vektor

$$\Theta = [l_1, l_2]'$$

Cílem měření je nalézt odhady $\hat{\beta}$ skutečných hodnot β a dále odhady parametrů $\hat{\Theta} = [l_1, l_2]'$ vystupujících v Langevinově funkci tak, aby platilo

$$\begin{aligned} g_i(\beta, \Theta) &= l_1 \cdot \coth(l_2 \cdot x_i) - \frac{l_1}{l_2 \cdot x_i} - \beta_i = 0, \\ i &= 1, \dots, N. \end{aligned}$$

2.1 Užití Levenberg–Marquardtova algoritmu

V dalších částech práce se budeme setkávat s problematikou nalezení počátečního řešení pro navržené algoritmy. Pro tento účel budeme využívat algoritmus Levenberg - Marquardt. Tento algoritmus se řadí do třídy nelineárních metod nejmenších čtverců, jejichž cílem je odhadnout hodnoty neznámých parametrů zkoumané nelineární funkce. Tento algoritmus je kombinací dvou numerických metod (metody největšího spádu a Gauss–Newtonovy metody).

V metodě největšího spádu jsou iterace dány vztahem

$$\Theta^{(k+1)} = \Theta^{(k)} - \lambda^k \nabla \mathbf{g}(\Theta^{(k)}), \quad (1)$$

kde

$$\lambda = \operatorname{argmin}_{\lambda} (\Theta^{(k)} - \lambda^k \nabla \mathbf{g}(\Theta^{(k)})).$$

Tato metoda užívá k výpočtu hodnot první derivace funkce \mathbf{g} .

Newtonova metoda využije i druhých derivací a iterace jsou určeny vzorcem

$$\Theta^{(k+1)} = \Theta^{(k)} - [\mathbf{H}(\Theta^{(k)})]^{-1} \nabla \mathbf{g}(\Theta^{(k)}). \quad (2)$$

kde

$$\mathbf{H}(\Theta^{(k)})$$

je Hessián, t.j. matice druhých derivací funkce \mathbf{g} v bodě $\Theta^{(k)}$.

Ve speciální optimalizační úloze – metodě nejmenších čtverců

$$\min \mathbf{G}(\Theta) = \|\mathbf{g}(\Theta)\|_2^2 = \mathbf{g}(\Theta)' \mathbf{g}(\Theta)$$

se používá Gauss–Newtonova metoda. Zde je Hessián v Gaussově metodě nahrazen násobením dvou Jacobiánů (matic prvních derivací) funkce \mathbf{g} , tzv. pseudo-hessiánu, který má tvar

$$[\mathbf{J}_g(\Theta^{(k)})' \mathbf{J}_g(\Theta^{(k)})].$$

Iterační předpis je

$$\Theta^{(k+1)} = \Theta^{(k)} - [\mathbf{J}_g(\Theta^{(k)})' \mathbf{J}_g(\Theta^{(k)})]^{-1} \mathbf{J}_g(\Theta^{(k)})' \mathbf{g}(\Theta^{(k)}). \quad (3)$$

nebo

$$[\mathbf{J}_g(\Theta^{(k)})' \mathbf{J}_g(\Theta^{(k)})] (\Delta \Theta)^{(k)} = -\mathbf{J}_g(\Theta^{(k)})' \mathbf{g}(\Theta^{(k)}). \quad (4)$$

Levenberg (1944) a Marquardt (1963) navrhli hybridní metodu mezi metodou největšího spádu a Gauss–Newtonovou metodou.

$$[\mathbf{J}_g(\Theta^{(k)})' \mathbf{J}_g(\Theta^{(k)}) + \lambda \mathbf{I}] (\Delta \Theta)^{(k)} = -\mathbf{J}_g(\Theta^{(k)})' \mathbf{g}(\Theta^{(k)}). \quad (5)$$

Pro $\lambda \rightarrow 0$ Levenberg–Marquardtova metoda přechází ke Gauss–Newtonově metodě a pro $\lambda \rightarrow \infty$ k metodě největšího spádu.

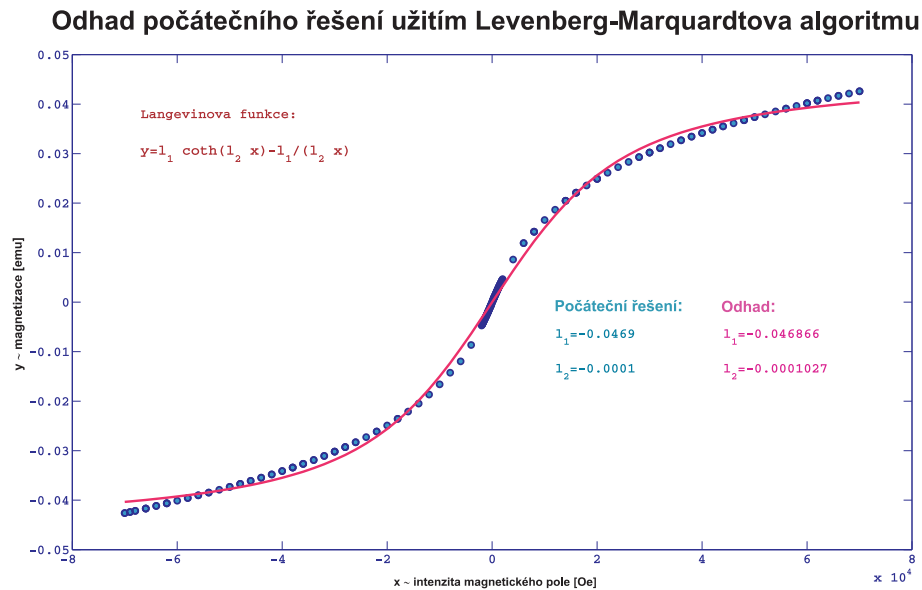
Pokud platí

$$\mathbf{g}(\Theta^{(k)} + (\Delta\Theta)^{(k)}) < \mathbf{g}(\Theta^{(k-1)} + (\Delta\Theta)^{(k-1)})$$

zmenšení λ (např. vydělením 2) vede ke zrychlení konvergence algoritmu. Nevýhodou tohoto numerického algoritmu je skutečnost, že není k dispozici odhad přesnosti nalezeného řešení.

Numerickou realizací Levenberg–Marquardtovy metody (obrázek 3.) dostáváme v naší úloze odhad počátečního řešení pro parametry l_1 a l_2 Langevinovy funkce:

$$\Theta_0 = \begin{pmatrix} -0,04686560 \\ -0,00010270 \end{pmatrix}$$



Obrázek 3: Odhad počátečního řešení

2.2 Design experimentu

Vzhledem k tomu, že měření probíhá ve 150 experimentálních bodech rovnoměrně, lze předpokládat, že některé z těchto bodů budou pro výsledné modelování

více či méně důležité. Proto se v této sekci budeme věnovat návrhu optimálního plánu měření za účelem zefektivnění měřícího procesu a zároveň zpřesnění odhadů výsledných parametrů. Z tohoto důvodu vytvoříme D - optimální plán měření, který minimalizuje konfidenční oblasti odhadovaných parametrů.

S využitím monografie [2], nadefinujeme postup, jenž vede k sestavení D - optimálního návrhu experimentu pro danou funkci.

Nejprve si vytvoříme známou matici plánu \mathbf{F} ve tvaru

$$\{\mathbf{F}\}_i = \frac{\partial \mathbf{f}(x_i, \Theta^0)}{\partial \Theta'} = \left(\frac{\partial \mathbf{f}(x_i, \Theta^0)}{\partial l_1}, \dots, \frac{\partial \mathbf{f}(x_i, \Theta^0)}{\partial l_k} \right), \quad (6)$$

kde

$$\mathbf{f}(x_i, \Theta) = l_1 \cdot \coth(l_2 \cdot x_i) - \frac{l_1}{l_2 \cdot x_i}.$$

i -tý řádek matice \mathbf{F} označíme \mathbf{f}'_i , jednoduchým výpočtem dostáváme

$$\mathbf{f}'_i = \left(\coth(l_2 \cdot x_i) - \frac{1}{l_2 \cdot x_i}, \frac{-l_1 \cdot x_i}{\sinh(l_2 \cdot x_i)^2} + \frac{l_1}{l_2^2 \cdot x_i} \right).$$

Definice 2.1. *Množina*

$$\epsilon = \{e_1, \dots, e_r\} \quad (7)$$

identifikačních znaků přímo měřitelných parametrů se nazývá množinou experimentálních bodů.

Definice 2.2. *Funkci $\delta : \{e_1, e_2, \dots, e_r\} \rightarrow \langle 0, 1 \rangle$ pro kterou platí*

$$\delta_i = \delta(e_i) \geq 0, \quad i = 1, 2, \dots, r, \quad \sum_{i=1}^r \delta(i) = 1, \quad (8)$$

budeme nazývat plán experimentu.

Číslo $\delta(i)$ udává relativní počet replikací i -té složky observačního vektoru.

Definice 2.3. *Spektrum návrhu experimentu je*

$$Sp(\delta) = \{e : \delta(e) > 0, e \in \epsilon\} \quad (9)$$

a obsahuje těch n ($n \leq r$) identifikačních znaků množiny experimentálních bodů, kterým plán δ přiřadil nenulovou hodnotu.

Definice 2.4. *Informační maticí experimentu nazveme*

$$\mathbf{M}(\delta) = \sum_{i \in Sp(\delta)} \delta(i) \lambda_i \mathbf{f}'_i, \quad (10)$$

kde \mathbf{f}'_i , $i = 1, \dots, r$ je i -tý řádek matice \mathbf{F} .

Popišme nyní způsob realizace měření pro dané $Sp(\delta)$.

Poznámka 2.1. Proky e_i z množiny $Sp(\delta)$ nám určují, že měření v i -tém bodě opakujeme r_i -krát. Platí $r_i = \delta(i)N$, kde N je celkový počet všech měření.

Věta 2.1. Nechť je dána realizace měření dle plánu experimentu uvedeného v předchozí poznámce. Pro tento návrh δ máme k dispozici nejlepší nestranný lineární odhad parametru (BLUE)

$$\widehat{\boldsymbol{\beta}}(\mathbf{Y}_\delta) = (\mathbf{F}'_\delta \boldsymbol{\Lambda}_\delta \mathbf{F}_\delta)^{-1} \mathbf{F}'_\delta \boldsymbol{\Lambda}_\delta \mathbf{Y}_\delta \quad (11)$$

\mathbf{F}_δ vytvořme z vybraných řádky \mathbf{f}_i , $i = 1, \dots, r$, pro které $e_i \in Sp(\delta)$ — tyto řádky označíme indexy i_1, \dots, i_n .

Obdobně matici $\boldsymbol{\Lambda}_\delta$ vytvořme z matice $\boldsymbol{\Lambda}$ vynecháním řádků a sloupců s indexy, kterým plán δ přiřadil nulovou hodnotu:

$$\boldsymbol{\Lambda}_\delta = \begin{pmatrix} \lambda_{i_1} r(i_1) & 0 & \dots & 0 \\ 0 & \lambda_{i_2} r(i_2) & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & \lambda_{i_n} r(i_n) \end{pmatrix}. \quad (12)$$

Pak kovarianční matice odhadu (11) je

$$\text{Var}(\widehat{\boldsymbol{\beta}}(\mathbf{Y}_\delta), N) = \frac{\sigma^2}{N} \mathbf{M}^{-1}(\delta) \quad (13)$$

Definice 2.5. Třídou všech návrhů δ takových, že $\mathbf{M}(\delta)$ je pozitivně definitní označíme $\boldsymbol{\Lambda}_{reg}$. Návrh δ_D^* je D -optimální, když

$$\det(\mathbf{M}^{-1}(\delta_D^*)) = \min \{ \det(\mathbf{M}^{-1}(\delta)), \delta \in \boldsymbol{\Lambda}_{reg} \}. \quad (14)$$

Věta 2.2. Následující tvrzení jsou ekvivalentní:

1. Návrh δ_D^* je D -optimální.

$$\begin{aligned} 2. \max \{ \lambda_i \mathbf{f}'_i \mathbf{M}^{-1}(\delta_D^*) \mathbf{f}_i : i = 1, \dots, r \} = \\ = \min \{ \max \{ \lambda_j \mathbf{f}'_j \mathbf{M}^{-1}(\delta_D^*) \mathbf{f}_j : j = 1, \dots, r \} : \delta \in \boldsymbol{\Lambda}_{reg} \}. \end{aligned} \quad (15)$$

$$\begin{aligned} 3. \max \{ \lambda_i \mathbf{f}'_i \mathbf{M}^{-1}(\delta_D^*) \mathbf{f}_i : i = 1, \dots, r \} = \\ = k. \end{aligned} \quad (16)$$

Důkaz: Viz [2]

Následující tvrzení poskytne iterační algoritmus pro nalezení optimálního návrhu.

Věta 2.3. *Nechť pro posloupnost návrhů $\delta_1, \delta_2, \dots$ platí*

$$\delta_{s+1} = (1 - \alpha^* s + 1)\delta_s + \alpha_{s+1}^* \lambda_{i_{s+1}^*} \mathbf{f}_{i_{s+1}^*} \mathbf{f}'_{i_{s+1}^*}, \quad (17)$$

kde i_{s+1}^* určíme z rovnice

$$\lambda_{i_{s+1}^*} d(e_{i_{s+1}^*}, \delta_s) = \max \{ \lambda_i d(e_i, \delta_s) : i = 1, \dots, r \}. \quad (18)$$

a kde λ_{s+1}^* určíme z rovnice

$$\alpha_{s+1}^* = \frac{\lambda_{i_{s+1}^*} d(e_{i_{s+1}^*}, \delta_s) - k}{[\lambda_{i_{s+1}^*} d(e_{i_{s+1}^*}, \delta_s) - 1] - k}, \quad (19)$$

kde

$$d(e_i, \delta) = \mathbf{f}'_i \mathbf{M}^{-1}(\delta_D^*) \mathbf{f}_i : \delta \in \mathbf{\Lambda}_{reg}. \quad (20)$$

Nechť $\delta_i \neq \delta_D^*$, $i = 1, 2, \dots$, pak

$$\lim_{n \rightarrow \infty} \det(\mathbf{M}(\delta_s)) = \det(\mathbf{M}(\delta_D^*)). \quad (21)$$

Důkaz: Viz [2]

Algoritmus 2.1. *Nechť měření můžeme provádět v experimentálních bodech $x = \{-70000, -68000, \dots, 68000, 70000\}$. Plán experimentu pro získání odhadů neznámých parametrů Langevinovy funkce má informační matici*

$$\mathbf{M}(\delta_i) = \begin{pmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{pmatrix}.$$

kde

$$\begin{aligned} M_{11} &= \sum_{i=1}^{150} \delta_i \cdot \left(\coth(l_2 \cdot x_i) - \frac{1}{l_2 \cdot x_i} \right)^2, \\ M_{12} &= \sum_{i=1}^{150} \delta_i \cdot \left(\left(\coth(l_2 \cdot x_i) - \frac{1}{l_2 \cdot x_i} \right) \left(\frac{-l_1 \cdot x_i}{\sinh(l_2 \cdot x_i)^2} + \frac{l_1}{l_2^2 \cdot x_i} \right) \right), \\ M_{22} &= \sum_{i=1}^{150} \delta_i \cdot \left(\frac{-l_1 \cdot x_i}{\sinh(l_2 \cdot x_i)^2} + \frac{l_1}{l_2^2 \cdot x_i} \right)^2. \end{aligned}$$

Chceme najít D -optimální plán, tedy najít index i^* pro který je výraz

$$\{ \lambda_i \mathbf{f}'_i \mathbf{M}^{-1}(\delta) \mathbf{f}_i : i = 1, \dots, r \}$$

maximální.

V bodě i^* zvýšíme počet měření a dále vytvoříme konvexní kombinaci naší informační matice s informační maticí jednobodového návrhu pro index i^* .

Uřídíme další iterace informační matice

$$\mathbf{M}(\delta_{i+1}) = (1 - \alpha)\mathbf{M}(\delta_i) + \alpha\lambda_{i^*}\mathbf{f}_{i^*}\mathbf{f}_{i^*}',$$

pro $i = 1, \dots, r$.

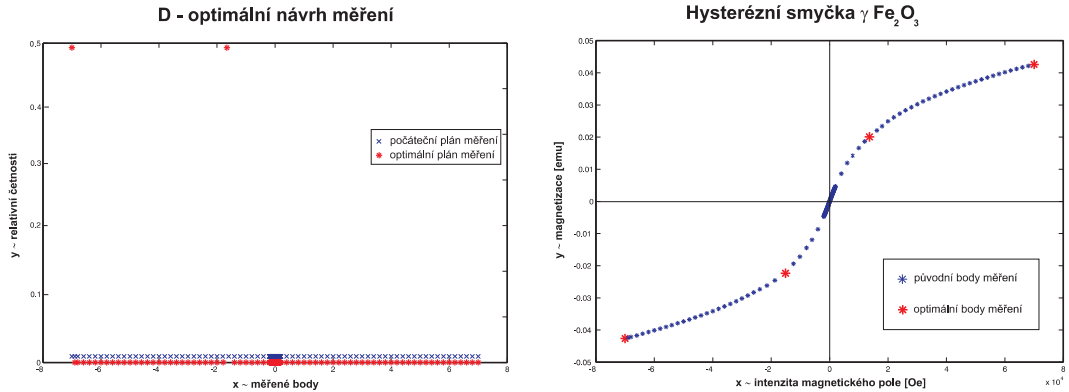
Algoritmus ukončíme pokud

$$\max \{ \lambda_i \mathbf{f}_i' \mathbf{M}^{-1}(\delta_D^*) \mathbf{f}_i : i = 1, \dots, r \} > k + \varepsilon.$$

Koeficient α před informační maticí jednobodového návrhu může být v prvním kroku převrácená hodnota počtu bodů ve spektru prvního návrhu. V dalších iteracích může být získáno převrácením čísla o jedna většího než v předcházející iteraci.

Poznámka 2.2. V uvedeném algoritmu je nutné v každém kroku realizovat výpočet nové inverzní matice. Zde lze užít trik, kdy novou inverzní matici určíme z předchozí. To sníží časovou složitost algoritmu. Matici $(\mathbf{M}(\delta_{i+1}))^{-1}$ lze dle Boxovy věty určit takto

$$\begin{aligned} (\mathbf{M}(\delta_{i+1}))^{-1} &= \frac{1}{1 - \alpha} (\mathbf{M}(\delta_i))^{-1} - \frac{1}{1 - \alpha} (\mathbf{M}(\delta_i))^{-1} \frac{\alpha \mathbf{f}_{i^*} \mathbf{f}_{i^*}'}{1 + \frac{\alpha}{1 - \alpha} \lambda_{i^*} \mathbf{f}_{i^*}' (\mathbf{M}(\delta_i))^{-1} \mathbf{f}_{i^*}} \\ &\quad \cdot \frac{1}{1 - \alpha} (\mathbf{M}(\delta_i))^{-1}. \end{aligned}$$



Obrázek 4: Optimální body měření Langevinovy funkce a jejich vizualizace na hysterézní smyčce

Na základě výše uvedeného postupu jsme získali body, jak je vidět z obrázku 4., v nichž je vhodné provést měření Langevinovy funkce. Tyto body můžeme považovat za optimální. Jedná se o měření v bodech -70000 Oe a -16000 Oe, přičemž relativní četnosti měření v těchto bodech jsou 0,498 a 0,499.

Vzhledem k tomu, že se jedná o plán měření symetrické funkce, je třeba ještě vydělit relativní četnosti měření v jednotlivých bodech dvěma a měření provést i v bodech 16000 Oe a 70000 Oe. Relativní četnosti měření v těchto čtyřech bodech budou tedy 0,249, 0,250, 0,250, 0,249.

3 Linearizované regresní modely

Z D - optimálního plánu měření Langevinovy funkce jsme získali množinu významných experimentálních bodů, pro které zkonstruujeme lineární regresní modely za účelem odhadu parametrů Langevinovy funkce. Nejprve se zaměříme na odhady parametrů v modelu bez podmínek a potom se zaměříme na model se systémem podmínek. Na závěr výsledky obou přístupů porovnáme.

3.1 Algoritmus 1 – Lineární regresní model nepřímého měření vektorového parametru bez podmínek

Provedené měření je ze statistického hlediska možné považovat za nelineární regresní model nepřímého měření vektorového parametru, ve tvaru

$$\mathbf{Y} \sim [\boldsymbol{\phi}(\boldsymbol{\Theta}), \boldsymbol{\Sigma}], \quad (22)$$

kde $\boldsymbol{\phi}(\boldsymbol{\Theta})$ je známá (v našem případě nelineární) funkce.

Abychom mohli přistoupit k výpočtu hodnot odhadů parametrů Langevinovy funkce, je potřebné provést linearizaci modelu (22). Je-li $\boldsymbol{\phi}(\boldsymbol{\Theta}^0)$ známý vektor, lze model rozvinout do Taylorovy řady v níž zanedbáme členy druhého a vyšších řádů. Po provedení linearizace lze model psát ve tvaru

$$\overline{\mathbf{Y}}_\delta \sim_n (\mathbf{F}_\delta \boldsymbol{\Theta}, \sigma^2 \boldsymbol{\Lambda}_\delta^{-1}), \quad (23)$$

kde $\boldsymbol{\Theta} = [l_1, l_2]'$ je vektor neznámých parametrů, $\sigma^2 \boldsymbol{\Lambda}_\delta^{-1}$ je varianční matice a

$$\{\mathbf{F}_\delta\}_i = \frac{\partial \phi_i(x_i, \boldsymbol{\Theta}^0)}{\partial \boldsymbol{\Theta}'} = \left(\frac{\partial \phi_i(x_i, \boldsymbol{\Theta}^0)}{\partial l_1}, \dots, \frac{\partial \phi_i(x_i, \boldsymbol{\Theta}^0)}{\partial l_k} \right), \quad (24)$$

přičemž

$$\begin{aligned} \frac{\partial \phi_i}{\partial l_1} &= \coth(l_2 \cdot x_i) - \frac{1}{l_2 \cdot x_i}, \\ \frac{\partial \phi_i}{\partial l_2} &= \frac{-l_1 \cdot x_i}{\sinh(l_2 \cdot x_i)^2} + \frac{l_1}{l_2^2 \cdot x_i} \end{aligned}$$

je známá matice plánu pro body x_i , kde $i \in \text{Sp}(\delta_D^*) = \{i : \delta_D^*(i) > 0\} = \{i_1, \dots, i_n\}$.
Nyní

$$\bar{\mathbf{Y}}_\delta = \begin{pmatrix} \frac{1}{r_{i_1}} \sum_{i=1}^{r_{i_1}} Y_{1,i} \\ \vdots \\ \frac{1}{r_{i_n}} \sum_{i=1}^{r_{i_n}} Y_{n,i} \end{pmatrix} \quad (25)$$

je tedy nový observační vektor s varianční maticí ve tvaru

$$\text{Var}(\bar{\mathbf{Y}}_\delta) = \sigma^2 \begin{pmatrix} \frac{1}{r_{i_1}} & & \\ & \ddots & \\ & & \frac{1}{r_{i_n}} \end{pmatrix} = \sigma^2 \mathbf{\Lambda}_\delta^{-1}. \quad (26)$$

Věta 3.1. *Nechť je dán linearizovaný regresní model nepřímého měření vektorového parametru (23). Pak BLUE parametru Θ je dán ve tvaru*

$$\hat{\Theta}_\delta = (\mathbf{F}'_\delta \mathbf{\Lambda}_\delta \mathbf{F}_\delta)^{-1} \mathbf{F}'_\delta \mathbf{\Lambda}_\delta \bar{\mathbf{Y}}_\delta, \quad (27)$$

s varianční maticí

$$\text{Var}(\hat{\Theta}_\delta) = \sigma^2 (\mathbf{F}'_\delta \mathbf{\Lambda}_\delta \mathbf{F}_\delta)^{-1}. \quad (28)$$

Důkaz: Viz [2]

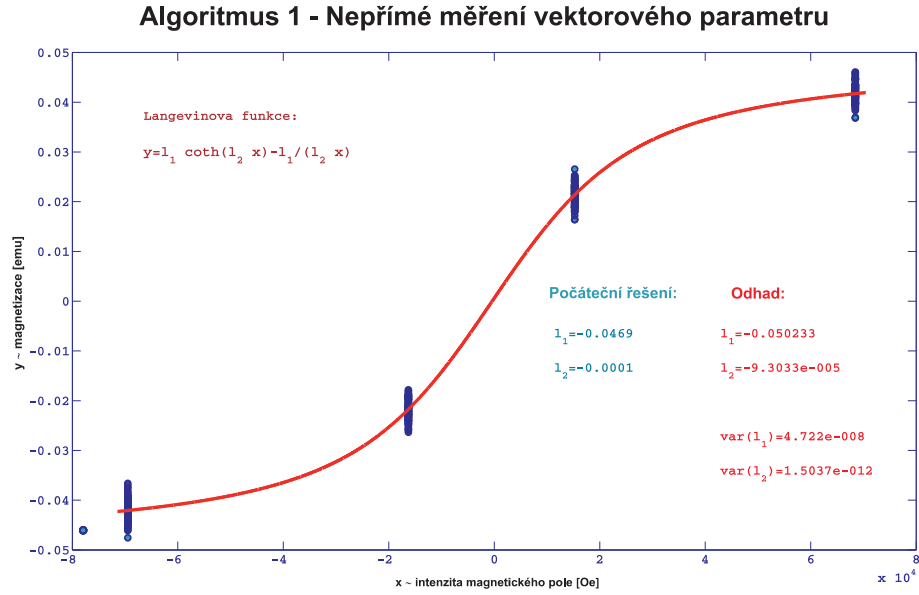
Po aplikaci výše uvedeného postupu jsme získali následující odhady parametrů l_1 a l_2 Langevinovy funkce

$$\hat{\Theta}_\delta = \begin{pmatrix} -0,05023335 \\ -0,00009303 \end{pmatrix}.$$

Variance těchto odhadů je potom

$$\text{Var}(\hat{\Theta}_\delta) = \begin{pmatrix} (2,1730 \cdot 10^{-4})^2 & -2,0673 \cdot 10^{-10} \\ -2,0673 \cdot 10^{-10} & (1,2262 \cdot 10^{-6})^2 \end{pmatrix}.$$

Grafická reprezentace výsledků je zobrazena na obrázku 5.



Obrázek 5: Model nepřímého měření vektorového parametru

3.2 Algoritmus 2 – Lineární regresní model přímého měření vektorového parametru se systémem podmínek typu II

Mějme hodnoty $(Y_1^*, Y_2^*, \dots, Y_N^*)'$ parametru β v bodech $(x_{i_1}, x_{i_2}, \dots, x_{i_n})'$ určených deterministicky. Hodnoty x_i , $i = i_1, \dots, i_n$ reprezentují stanovenou sílu vnějšího magnetického pole a hodnoty Y_i^* , $i = 1, \dots, N$ představují měření odpovídající magnetizace materiálu v bodech vybraných D - optimálním návrhem měření, přičemž celkový počet měření (N) zůstal zachován. Hodnoty závislé proměnné mohou být vyjádřeny pomocí Langevinovy funkce takto:

$$Y_i^* = l_1 \cdot \coth(l_2 \cdot x_i) - \frac{l_1}{l_2 \cdot x_i}, \quad i = i_1, \dots, i_n. \quad (29)$$

Nelineární model měření lze tedy popsat

$$\overline{Y}_\delta \sim N_n \left[\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}, \Sigma \right], \quad (30)$$

příčemž musí platit

$$g_i(\boldsymbol{\beta}, \boldsymbol{\Theta}) = l_1 \cdot \coth(l_2 \cdot x_i) - \frac{l_1}{l_2 \cdot x_i} - \beta_i = 0, \quad (31)$$

$$i = i_1, \dots, i_n,$$

kde

$$\boldsymbol{\Theta} = [l_1, l_2]',$$

je vektor neznámých parametrů.

Cílem měření je nalézt odhady $\widehat{\boldsymbol{\beta}}_\delta$ skutečných hodnot $\boldsymbol{\beta}$ a dále odhady parametrů $\widehat{\boldsymbol{\Theta}}_\delta = [l_1, l_2]'$ vystupujících v Langevinově funkci tak, aby splnili (32).

Stejně, jako tomu bylo u modelu nepřímého měření vektorového parametru bez podmínek, je potřebné i zde provést linearizaci modelu pomocí Taylorova rozvoje.

Nelineární podmínky $\mathbf{g}(\boldsymbol{\beta}, \boldsymbol{\Theta}) = (g_1(\boldsymbol{\beta}, \boldsymbol{\Theta}), \dots, g_n(\boldsymbol{\beta}, \boldsymbol{\Theta}))' = 0$ lze po linearizaci pomocí Taylorova rozvoje psát v symbolickém zápise $\mathbf{B}\delta\boldsymbol{\beta} + \mathbf{G}\delta\boldsymbol{\Theta} + \mathbf{b} = \mathbf{0}$, kde $\mathbf{B} = \frac{\partial \mathbf{g}(\boldsymbol{\beta}^0, \boldsymbol{\Theta}^0)}{\partial \boldsymbol{\beta}'}$, $\mathbf{G} = \frac{\partial \mathbf{g}(\boldsymbol{\beta}^0, \boldsymbol{\Theta}^0)}{\partial \boldsymbol{\Theta}'}$, a $\mathbf{b} = \mathbf{g}(\boldsymbol{\beta}^0, \boldsymbol{\Theta}^0)$ v přibližném řešení $(\boldsymbol{\beta}^0, \boldsymbol{\Theta}^0)$.

Dále matice \mathbf{B} má strukturu

$$\mathbf{B} = \begin{pmatrix} b_{11} & 0 & \dots & 0 & 0 \\ 0 & b_{22} & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & 0 & b_{nn} \end{pmatrix}, \quad (32)$$

kde

$$b_{ii} = -1$$

$$\mathbf{G} = \frac{\partial \mathbf{g}(\boldsymbol{\beta}^0, \boldsymbol{\Theta}^0)}{\partial \boldsymbol{\Theta}'} = \begin{pmatrix} \frac{\partial g_1}{\partial l_1} & \frac{\partial g_1}{\partial l_2} \\ \vdots & \vdots \\ \frac{\partial g_i}{\partial l_1} & \frac{\partial g_i}{\partial l_2} \\ \vdots & \vdots \\ \frac{\partial g_n}{\partial l_1} & \frac{\partial g_n}{\partial l_2} \end{pmatrix} \quad (33)$$

kde

$$\frac{\partial g_i}{\partial l_1} = \coth(l_2 \cdot x_i) - \frac{1}{l_2 \cdot x_i},$$

$$\frac{\partial g_i}{\partial l_2} = \frac{-l_1 \cdot x_i}{\sinh(l_2 \cdot x_i)^2} + \frac{l_1}{l_2^2 \cdot x_i},$$

kde $i = i_1, \dots, i_n$

Nyní najdeme odhad parametrů v linearizovaném modelu. K tomuto účelu využijeme teorii lineárních statistických modelů uvedenou v [2], resp. model přímého měření vektorového parametru se systémem podmínek typu II.

Definice 3.1. *Model přímého měření s podmínkou II. typu na parametry 1. řádu má tvar*

$$\mathbf{Y}_\delta \sim_n (\mathbf{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{\Lambda}_\delta^{-1}), \quad (34)$$

$$\mathbf{b} + \mathbf{B}\boldsymbol{\beta} + \mathbf{G}\boldsymbol{\Theta} = \mathbf{0}, \quad (35)$$

kde $\boldsymbol{\beta} \in R^{k_1}$, $\boldsymbol{\Theta} \in R^{k_2}$ jsou neznámé.

Jestliže $h(\mathbf{X}_{(n,k)}) = k_1 < n$, $h(\mathbf{B}_{(q,k_1)}, \mathbf{G}_{(q,k_2)}) = q < k_1 + k_2$, $h(\mathbf{G}) = k_2 < q$ a matice $\boldsymbol{\Lambda}_\delta^{-1}$ je pozitivně definitní, potom model nazýváme regulárním.

Dále budeme uvažovat pouze regulární model.

Věta 3.2. *BLUE vektoru $\begin{pmatrix} \widehat{\boldsymbol{\beta}}_\delta \\ \widehat{\boldsymbol{\Theta}}_\delta \end{pmatrix}$ jsou dány vztahy*

$$\widehat{\boldsymbol{\beta}}_\delta = \widehat{\boldsymbol{\beta}} - (\mathbf{X}'\boldsymbol{\Lambda}_\delta\mathbf{X})^{-1}\mathbf{B}' [\mathbf{T}^{-1} - \mathbf{T}^{-1}\mathbf{G}(\mathbf{G}'\mathbf{T}^{-1}\mathbf{G})^{-1}\mathbf{G}'\mathbf{T}^{-1}] (\mathbf{b} + \mathbf{B}\widehat{\boldsymbol{\beta}}) \quad (36)$$

$$\widehat{\boldsymbol{\Theta}}_\delta = -(\mathbf{G}'\mathbf{T}^{-1}\mathbf{G})^{-1}\mathbf{G}'\mathbf{T}^{-1}(\mathbf{b} + \mathbf{B}\widehat{\boldsymbol{\beta}}) \quad (37)$$

kde

$$\mathbf{T} = \mathbf{B}(\mathbf{X}'\boldsymbol{\Lambda}_\delta\mathbf{X})^{-1}\mathbf{B}' + \mathbf{G}\mathbf{G}', \quad (38)$$

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\boldsymbol{\Lambda}_\delta\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Lambda}_\delta\mathbf{Y}_\delta \quad (39)$$

($\widehat{\boldsymbol{\beta}}_\delta$ je odhad nerespektující podmínku týkající se parametrů $\boldsymbol{\beta}, \boldsymbol{\Theta}$).

Důkaz: Viz [4]

Věta 3.3. *Kovarianční matice odhadu $\widehat{\boldsymbol{\beta}}_\delta$ je*

$$\begin{aligned} \text{Var}(\widehat{\boldsymbol{\beta}}_\delta) = & \sigma^2 \{ \mathbf{I} - (\mathbf{X}'\boldsymbol{\Lambda}_\delta\mathbf{X})^{-1}\mathbf{B}' [\mathbf{T}^{-1} - \mathbf{T}^{-1}\mathbf{G}(\mathbf{G}'\mathbf{T}^{-1}\mathbf{G})^{-1}\mathbf{G}'\mathbf{T}^{-1}] \mathbf{B} \} \cdot \\ & (\mathbf{X}'\boldsymbol{\Lambda}_\delta\mathbf{X})^{-1} \{ \mathbf{I} - (\mathbf{X}'\boldsymbol{\Lambda}_\delta\mathbf{X})^{-1}\mathbf{B}' \cdot \\ & \cdot [\mathbf{T}^{-1} - \mathbf{T}^{-1}\mathbf{G}(\mathbf{G}'\mathbf{T}^{-1}\mathbf{G})^{-1}\mathbf{G}'\mathbf{T}^{-1}] \mathbf{B}' \} . \end{aligned} \quad (40)$$

Důkaz: Viz [4]

Věta 3.4. Kovarianční matice odhadu $\widehat{\Theta}_\delta$ je

$$\text{Var}(\widehat{\Theta}_\delta) = \sigma^2 \{(\mathbf{G}'\mathbf{T}^{-1}\mathbf{G})^{-1} - \mathbf{I}\} . \quad (41)$$

Důkaz: Viz [4]

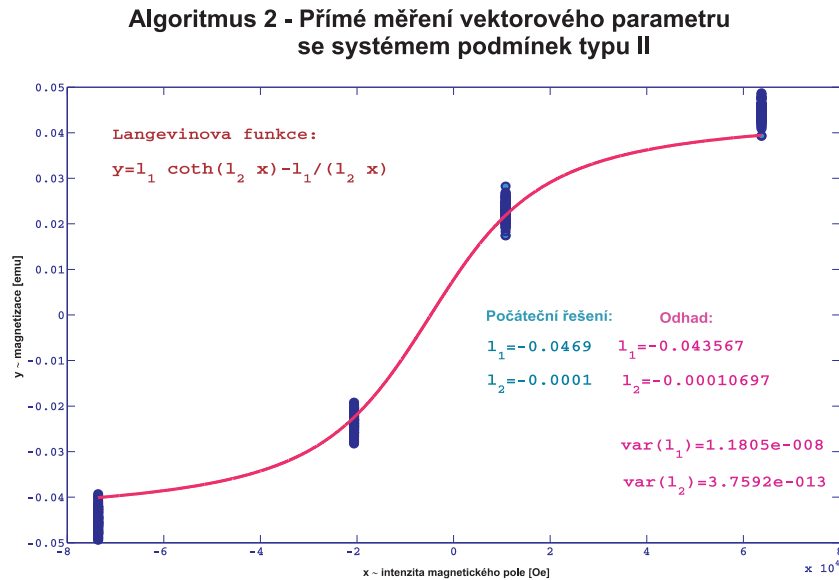
Na obrázku 6. jsou uvedeny výsledky dosažené na základě výše uvedeného modelu. Naším úkolem bylo opět odhadnout parametry Langevinovy funkce. Jak již bylo řečeno, k odhadu počátečního řešení jsme použili Levenberg - Marquardtův algoritmus.

Odhady parametrů l_1 a l_2 na základě měření provedeného dle D - optimálního plánu jsou následující:

$$\widehat{\Theta}_\delta = \begin{pmatrix} -0,04356664 \\ -0,00010697 \end{pmatrix}$$

a variace těchto odhadů je potom

$$\text{Var} \left(\widehat{\Theta}_\delta \right) = \begin{pmatrix} (1,0865 \cdot 10^{-4})^2 & -5,1684 \cdot 10^{-11} \\ -5,1684 \cdot 10^{-11} & (6,1312 \cdot 10^{-7})^2 \end{pmatrix} .$$



Obrázek 6: Model přímého měření vektorového parametru s podmínkami typu II

3.3 Porovnání vhodnosti volby modelů

Abychom mohli porovnat kvalitu obou výše zmíněných nelineárních regresních modelů, použijeme jako kritérium tzv. reziduální součet čtverců, který je definovaný jako

$$S = \sum_{i=1}^n (Y_{\delta_i} - \widehat{Y}_{\delta_i}), \quad (42)$$

kde Y_{δ_i} a \widehat{Y}_{δ_i} jsou pozorované, resp. vyrovnané hodnoty závisle proměnné \mathbf{Y}_δ . Podle tohoto kritéria je vhodnější ten model, pro nějž nabývá tato statistika nižší hodnoty.

Protože počet regresních parametrů významně ovlivňuje výslednou hodnotu S , je vhodnější pro srovnání použít tzv. reziduální rozptyl, tedy

$$s_e^2 = \frac{S}{n - p}, \quad (43)$$

kde n je počet uskutečněných měření a p je počet regresních parametrů.

Jako další kritérium pro srovnání modelů jsme použili tzv. index determinace definovaný jako

$$I^2 = 1 - \frac{\sum_{i=1}^n (\widehat{Y}_{\delta_i} - Y_{\delta_i})^2}{\sum_{i=1}^n (Y_{\delta_i} - \overline{Y}_{\delta_i})^2} \quad (44)$$

a který nabývá hodnot z intervalu $\langle 0, 1 \rangle$.

Všechny dosažené výsledky a tato srovnání uvádíme v přehledové tabulce (1).

Parametr	Model bez podmínky	Model se systémem podm. typu II
Odhad l_1	-0,05023335	-0,04356664
Odhad l_2	-0,00009303	-0,00010697
Var (l_1)	$(2,1730 \cdot 10^{-4})^2$	$(1,0865 \cdot 10^{-4})^2$
Var (l_2)	$(1,2262 \cdot 10^{-6})^2$	$(6,1312 \cdot 10^{-7})^2$
S	$1,8478 \cdot 10^{-3}$	$7,3394 \cdot 10^{-3}$
s_e^2	$4,1247 \cdot 10^{-6}$	$1,6383 \cdot 10^{-5}$
Index determinace	0,995626	0,993818

Tabulka 1: Srovnání hodnot pro odhady parametrů Langevinovy funkce pomocí použitých regresních modelů

Na základě těchto výsledků můžeme konstatovat, že pro odhad neznámých parametrů Langevinovy funkce je vhodnější použít model bez podmínky.

4 Závěr

V předkládané práci jsme se zabývali problematikou odhadu neznámých parametrů Langevinovy funkce. Za tímto účelem jsme danou funkci linearizovali a sestavili jsme dva lineární modely. Model nepřímého měření vektorového parametru bez podmínek a model přímého měření vektorového parametru se systémem podmínek typu II. U původního plánu experimentu byl tedy observační vektor tvořen třemi opakovanými měřeními, jež byli provedeny ve 150 bodech vybraných z intervalu -70000 Oe až 70000 Oe.

Kovarianční matice odhadů těchto neznámých parametrů byli u obou modelů téměř totožné, tedy

$$\text{Var}(\hat{\Theta}) = \text{Var}(\hat{\Theta}) = \begin{pmatrix} (4,2640 \cdot 10^{-4})^2 & -9,2596 \cdot 10^{-10} \\ -9,2596 \cdot 10^{-10} & (2,4668 \cdot 10^{-6})^2 \end{pmatrix}.$$

Shoda výsledků při použití dvou různých lineárních modelů signalizuje, že proces linearizace ovlivnil výsledky pouze nepatrně.

Cílem této práce však bylo především optimalizovat experimentální měření. Abychom našli body, které jsou skutečně významné pro experimentální měření, navrhli jsme D - optimální plán měření. Výsledky tohoto přístupu jsme dále použili pro sestavení již výše zmíněných lineárních modelů. V tomto případě už byl však observační vektor odlišný. 450 měření, jež jsme měli k dispozici, jsme rozložili pouze do těch bodů, jež byly D - optimálním plánem vyhodnoceny jako významné. Z optimalizovaného plánu experimentu jsme dosáhli těchto výsledků.

Získali jsme kovarianční matici odhadů získaných z modelu nepřímého měření vektorového parametru bez podmínek

$$\text{Var}(\hat{\Theta}_\delta) = \begin{pmatrix} (2,1730 \cdot 10^{-4})^2 & -2,0673 \cdot 10^{-10} \\ -2,0673 \cdot 10^{-10} & (1,2262 \cdot 10^{-6})^2 \end{pmatrix}$$

a kovarianční matici odhadů získanou z modelu přímého měření vektorového parametru se systémem podmínek typu II

$$\text{Var}(\hat{\Theta}_\delta) = \begin{pmatrix} (1,0865 \cdot 10^{-4})^2 & -5,1684 \cdot 10^{-11} \\ -5,1684 \cdot 10^{-11} & (6,1312 \cdot 10^{-7})^2 \end{pmatrix}.$$

Porovnání kovariančních matic z původního plánu experimentu a z optimalizovaného plánu experimentu ukazuje nečekaně výrazné vylepšení přesnosti odhadů. Za úspěch lze považovat i takový výsledek, který by znamenal 10-ti nebo 20-ti procentí zmenšení směrodatných odchylek. Efekt optimalizace je však v našem případě daleko výraznější.

Literatura

- [1] Jiří Anděl: Základy matematické statistiky, Univerzita Karlova v Praze, Praha 2002
- [2] Lubomír Kubáček, Ludmila Kubáčková: Statistika a metrologie, UP Olomouc, 2000
- [3] Jaroslav Marek: Estimation in connecting measurements, Acta, UP Olomouc, 2004
- [4] Jaroslav Marek, Pavel Tuček: Statistické algoritmy pro aproximaci Lorentzovy funkce, KMAaAM, Preprint 9/2009