

TECHNICKÁ UNIVERZITA V LIBERCI

Fakulta mechatroniky a mezioborových inženýrských studií

Ing. Radomír Kamler

**METODY IDENTIFIKACE SPOJITÝCH MODELŮ
TECHNOLOGICKÝCH SOUSTAV**

Ing. Radomír Kamler

Školitel: Doc. Ing. Bedřich Janeček, CSc.

Disertační práce

Liberec 1999

U343-02

TECHNICKÁ UNIVERZITA V LIBERCI

Fakulta mechatroniky a mezioborových inženýrských studií

Obor: Technická kybernetika

Katedra řídicí techniky

METODY IDENTIFIKACE SPOJITÝCH MODELŮ TECHNOLOGICKÝCH SOUSTAV

Ing. Radomír Kamler

Školitel: Doc. Ing. Bedřich Janeček, CSc.

UNIVERZITNÍ KNIHOVNA
TECHNICKÉ UNIVERZITY V LIBERCI



3146070696

Rozsah práce a příloh:

Počet stran.....	153
Počet tabulek.....	11
Počet obrázků.....	91

Místopřísežné prohlášení:

Místopřísežně prohlašuji, že jsem disertační práci vypracoval samostatně s použitím uvedené literatury.

V Liberci 29. 6. 1999

.....
Ing. Radomír Kamler

Práce popsaná na následujících stranách by nikdy nemohla vzniknout bez pomoci několika osobností, kterým bych zde rád poděkoval.

Jsem velice zavázán svým školitelům, Doc. Ing. Bedřichu Janečkovi, Csc. a Doc. Ing. Osvaldu Modrlákovi, Csc., kteří mě uvedli do problematiky identifikace technologických procesů a v průběhu vzniku této práce mi pomáhali cennými diskusemi, připomínkami a podněty. Dále bych také velice rád poděkoval svému školiteli v průběhu mé stáže na Univerzitě v Cincinnati Prof. Garymu L. Slaterovi Ph.D. za odborné vedení a veškerou pomoc při uskutečnění tohoto studijního pobytu.

A v neposlední řadě nemohu opomenout poděkovat ostatním členům katedry řídicí techniky a všem, kteří se jakkoli podíleli na vzniku této práce za jejich vynaložené úsilí spojené s mým doktorandským studiem.

.....
Radomír Kamler

OBSAH

SEZNAM OBRÁZKŮ	4
SEZNAM TABULEK.....	7
SEZNAM ZÁKLADNÍCH SYMBOLŮ.....	8
ABSTRACT.....	9
I ÚVODNÍ POZNÁMKA.....	10
I.1 CÍLE DISERTAČNÍ PRÁCE	11
II EXPERIMENTÁLNÍ IDENTIFIKACE	12
II.1 ÚVOD.....	13
II.1.1 ÚČEL MODELOVÁNÍ	13
II.1.2 SYSTÉM	13
II.1.3 MODEL.....	14
II.1.4 FIKCE PRAVDIVÉHO MODELU.....	15
II.1.5 POSTUP IDENTIFIKACE	15
II.1.5.1 Data, modely, kritéria	15
II.1.5.2 Ověření platnosti modelu.....	17
II.1.5.3 Identifikační cyklus	17
II.1.6 IDENTIFIKACE VE SPOJITÉ ČASOVÉ OBLASTI - SOUČASNÝ STAV	19
II.2 STRUKTURY MODELŮ.....	20
II.2.1 FYZIKÁLNÍ VERSUS „BLACK BOX“ MODEL Y.....	20
II.2.2 LINEÁRNÍ VERSUS NELINEÁRNÍ MODEL Y.....	22
II.2.3 MODEL Y VE SPOJITÉ VERSUS V DISKRÉTNÍ OBLASTI.....	23
II.2.4 DETERMINISTICKÉ VERSUS STOCHASTICKÉ MODEL Y	25
II.2.5 ZÁVĚR	28
II.3 OPTIMALIZACE.....	29
II.3.1 KRITÉRIUM JAKOSTI MODEL U	29

II.3.1.1	Kritérium nejmenších čtverců.....	30
II.3.1.2	Kritérium absolutních hodnot.....	31
II.3.1.3	Kritérium maximální věrohodnosti	31
II.3.2	METODA NEJMENŠÍCH ČTVERCŮ.....	32
II.3.2.1	Numerické řešení metody nejm. čtverců	32
II.3.3	REKURZIVNÍ METODA NEJMENŠÍCH ČTVERCŮ	33
II.3.3.1	Vektor p se předpokládá konstantní.....	34
II.3.3.2	Vektor p může kolísat	36
II.3.4	OBEČNÉ METODY OPTIMALIZACE	37
II.3.4.1	Částečně lineární model	38
II.3.4.2	Jednodimenzionální optimalizace	39
II.3.4.2.1	Volba prohledávaného intervalu	39
II.3.4.2.2	Dichotomické vyhledávání.....	40
II.3.4.2.3	Fibonacciho metoda a metoda zlatého řezu	41
II.3.4.2.4	Parabolická interpolace	43
II.3.4.3	Optimalizace pomocí prvních a druhých derivací účelové funkce	43
II.3.4.3.1	Metoda gradientu	43
II.3.4.3.2	Newtonova metoda	44
II.3.4.3.3	Metoda konjugovaných gradientů.....	45
II.3.4.4	Globální optimalizace.....	47
II.3.4.4.1	Náhodné prohledávání	47
II.4	VYHODNOCENÍ KVALITY MODELU	49
III	IDENTIFIKACE VE SPOJITÉ ČASOVÉ OBLASTI	51
III.1	ÚVOD	52
III.2	INTEGRAČNÍ METODA	53
III.3	DERIVAČNÍ METODA	58
III.4	VLIV POČÁTEČNÍCH PODMÍNEK	64
III.4.1	APLIKACE NENULOVÝCH POČÁTEČNÍCH PODMÍNEK NA INTEGRAČNÍ METODU	65
III.4.2	APLIKACE NENULOVÝCH POČÁTEČNÍCH PODMÍNEK NA DERIVAČNÍ METODU	67
III.5	INTEGRAČNÍ A DERIVAČNÍ METODA PRO NELINEÁRNÍ SYSTÉMY	70
III.6	METODA KVAZILINEARIZACE	72
III.7	KOMBINACE KVAZILINEARIZACE A INTEGRAČNÍ METODY	75
III.8	IDENTIFIKACE NEDIFERENCovatELNÝCH NELINEÁRNÍCH MODELŮ	78

III.9	ROZBOR VLIVU PORUCH NA ODHAD PARAMETRŮ.....	83
III.10	ZÁVĚR	94
IV	ŘEŠENÉ ÚLOHY	95
IV.1	IDENTIFIKACE PRŮTOKOVÉHO VÝMĚNÍKU	96
IV.2	IDENTIFIKACE STEJNOSMĚRNÉHO MOTORKU.....	105
IV.3	IDENTIFIKACE SS MOTORKU S PRUŽNOU SPOJKOU	118
IV.4	IDENTIFIKACE VÝŠKY HLADINY	132
V	ZÁVĚR.....	142
	SEZNAM LITERATURY.....	144
	REJSTŘÍK	148

SEZNAM OBRÁZKŮ

OBRÁZEK II.1-1: SYSTÉM.....	14
OBRÁZEK II.1-2: MODEL	14
OBRÁZEK II.1-3: INVERZNÍ MODEL	15
OBRÁZEK II.1-4: TOK INFORMACÍ PŘI OPTIMALIZACI	16
OBRÁZEK II.1-5: IDENTIFIKAČNÍ CYKLUS.....	18
OBRÁZEK II.2-1: STRUKTURY MODELŮ.....	28
OBRÁZEK II.3-1: DICHOTOMICKÉ VYHLEDÁVÁNÍ	40
OBRÁZEK II.3-2: PRINCIP FIBONACIHO METODY A METODY ZLATÉHO ŘEZU	41
OBRÁZEK II.3-3:ILUSTRACE DĚLENÍ INTERVALU, A. FIBONACIHO METODA, B. METODA ZLATÉHO ŘEZU.	42
OBRÁZEK II.3-4: KONJUGOVANÉ SMĚRY HLEDÁNÍ	46
OBRÁZEK III.2-1: LINEÁRNÍ JEDNOROZMĚRNÝ MODEL	53
OBRÁZEK III.2-2: POROVNÁNÍ VÝSTUPŮ MODELU A SYSTÉMU	56
OBRÁZEK III.2-3: ÚČELOVÁ FUNKCE, A. $P(3)=1$, $P(1),P(2)$ SE MĚNÍ, B. $P(2)=1$, $P(1),P(3)$ SE MĚNÍ.....	57
OBRÁZEK III.3-1: LINEÁRNÍ JEDNOROZMĚRNÝ MODEL	58
OBRÁZEK III.3-2: MODULAČNÍ FUNKCE A JEJÍ DERIVACE (ZE ZDOLA: $\phi(T)$, $\phi'(T)$, $\phi''(T)$, $\phi'''(T)$).....	61
OBRÁZEK III.3-3: PŘÍKLAD MODULAČNÍ FUNKCE A JEJÍCH DERIVACÍ (ZE ZDOLA: $\phi(T)$, $\phi'(T)$, $\phi''(T)$)	62
OBRÁZEK III.3-4: POROVNÁNÍ VÝSTUPŮ MODELU A SYSTÉMU	63
OBRÁZEK III.4-1:ZMĚŘENÝ VÝSTUP	66
OBRÁZEK III.4-2: POROVNÁNÍ VÝSTUPŮ MODELU A SYSTÉMU	67
OBRÁZEK III.4-3: POROVNÁNÍ VÝSTUPŮ MODELU A SYSTÉMU	69
OBRÁZEK III.5-1: ZMĚŘENÁ VSTUPNĚ-VÝSTUPNÍ CHARAKTERISTIKA.....	71
OBRÁZEK III.5-2: POROVNÁNÍ VÝSTUPŮ MODELU A SYSTÉMU	71
OBRÁZEK III.6-1: POROVNÁNÍ VÝSTUPŮ MODELU A SYSTÉMU	74
OBRÁZEK III.6-2: PRŮBĚH ÚČELOVÉ FUNKCE	74
OBRÁZEK III.7-1: ZMĚŘENÁ VSTUPNĚ-VÝSTUPNÍ CHARAKTERISTIKA.....	76
OBRÁZEK III.7-2: POROVNÁNÍ VÝSTUPŮ MODELU A SYSTÉMU	77
OBRÁZEK III.8-1: TYPICKÉ PŘÍKLADY TVRDÝCH NELINEARIT.....	78
OBRÁZEK III.8-2: ZMĚŘENÁ VSTUPNĚ-VÝSTUPNÍ CHARAKTERISTIKA.....	81
OBRÁZEK III.8-3: POROVNÁNÍ VÝSTUPŮ MODELU A SYSTÉMU	82
OBRÁZEK III.9-1: PORUCHY VSTUPUJÍCÍ DO SYSTÉMU.....	84
OBRÁZEK III.9-2: KVADRATICKÁ CHYBA J_1 V ZÁVISLOSTI NA ÚROVNI PORUCHY PŮSOBÍCÍ NA VÝSTUPU	87
OBRÁZEK III.9-3: KVADRATICKÁ CHYBA J_2 V ZÁVISLOSTI NA ÚROVNI PORUCHY PŮSOBÍCÍ NA VÝSTUPU	88
OBRÁZEK III.9-4: KVADRATICKÁ CHYBA J_1 V ZÁVISLOSTI NA ÚROVNI PORUCHY PŮSOBÍCÍ NA VSTUPU	89

OBRÁZEK III.9-5: KVADRATICKÁ CHYBA J_2 V ZÁVISLOSTI NA ÚROVNI PORUCHY PŮSOBÍCÍ NA VSTUPU	90
OBRÁZEK III.9-6: KVADRATICKÁ CHYBA J_1 V ZÁVISLOSTI NA ÚROVNI PORUCHY VE TVARU DRIFTU PŮSOBÍCÍ NA VÝSTUPU	91
OBRÁZEK III.9-7: KVADRATICKÁ CHYBA J_2 V ZÁVISLOSTI NA ÚROVNI PORUCHY VE TVARU DRIFTU PŮSOBÍCÍ NA VÝSTUPU	91
OBRÁZEK III.9-8: KVADRATICKÁ CHYBA J_1 V ZÁVISLOSTI NA ÚROVNI PORUCHY VE TVARU DRIFTU PŮSOBÍCÍ NA VSTUPU	92
OBRÁZEK III.9-9: KVADRATICKÁ CHYBA J_2 V ZÁVISLOSTI NA ÚROVNI PORUCHY VE TVARU DRIFTU PŮSOBÍCÍ NA VSTUPU	92
OBRÁZEK III.9-10: KVADRATICKÁ CHYBA J_1 V ZÁVISLOSTI NA VELIKOSTI VZORKOVACÍ PERIODY PRO INTEGRAČNÍ METODU	93
OBRÁZEK III.9-11: KVADRATICKÁ CHYBA J_1 V ZÁVISLOSTI NA VELIKOSTI VZORKOVACÍ PERIODY PRO DERIVAČNÍ METODU	93
OBRÁZEK IV.1-1: PRŮTOKOVÝ VÝMĚNÍK - SCHÉMA SYSTÉMU	96
OBRÁZEK IV.1-2: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU VÝMĚNÍKU NA DATECH POUŽITÝCH PRO OPTIMALIZACI	99
OBRÁZEK IV.1-3: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU VÝMĚNÍKU NA JINÝCH DATECH NEŽ BYLA POUŽITA PRO OPTIMALIZACI	99
OBRÁZEK IV.1-4: ČASOVÁ POSLOUPNOST RESIDUÍ MODELU VÝMĚNÍKU	100
OBRÁZEK IV.1-5: GRAF RESIDUÍ V ZÁVISLOSTI NA BUDÍCÍM SIGNÁLU	100
OBRÁZEK IV.1-6: AUTOKOVARIANČNÍ FUNKCE RESIDUÍ MODELU VÝMĚNÍKU	101
OBRÁZEK IV.1-7: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU VÝMĚNÍKU NA DATECH POUŽITÝCH PRO OPTIMALIZACI	102
OBRÁZEK IV.1-8: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU VÝMĚNÍKU NA JINÝCH DATECH NEŽ BYLA POUŽITA PRO OPTIMALIZACI	102
OBRÁZEK IV.1-9: ČASOVÁ POSLOUPNOST RESIDUÍ MODELU VÝMĚNÍKU	103
OBRÁZEK IV.1-10: GRAF RESIDUÍ V ZÁVISLOSTI NA BUDÍCÍM SIGNÁLU	103
OBRÁZEK IV.1-11: AUTOKOVARIANČNÍ FUNKCE RESIDUÍ MODELU VÝMĚNÍKU	104
OBRÁZEK IV.2-1: STEJNOSMĚRNÝ MOTOR - SCHÉMA SYSTÉMU	105
OBRÁZEK IV.2-2: STEJNOSMĚRNÝ MOTOR - PRINCIPIÁLNÍ SCHÉMA	106
OBRÁZEK IV.2-3: COULOMBOVO A VISKÓZNÍ TŘENÍ	106
OBRÁZEK IV.2-4: STATICKÁ CHARAKTERISTIKA STEJNOSMĚRNÉHO MOTORKU	108
OBRÁZEK IV.2-5: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU SS MOTORKU NA DATECH POUŽITÝCH PRO OPTIMALIZACI	110
OBRÁZEK IV.2-6: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU SS MOTORKU NA JINÝCH DATECH NEŽ BYLA POUŽITA PRO OPTIMALIZACI	110
OBRÁZEK IV.2-7: ČASOVÁ POSLOUPNOST RESIDUÍ MODELU SS MOTORKU	111
OBRÁZEK IV.2-8: GRAF RESIDUÍ V ZÁVISLOSTI NA BUDÍCÍM SIGNÁLU	111
OBRÁZEK IV.2-9: AUTOKOVARIANČNÍ FUNKCE RESIDUÍ MODELU SS MOTORKU	112
OBRÁZEK IV.2-10: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU SS MOTORKU NA DATECH POUŽITÝCH PRO OPTIMALIZACI	114
OBRÁZEK IV.2-11: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU SS MOTORKU NA JINÝCH DATECH NEŽ BYLA POUŽITA PRO OPTIMALIZACI	115
OBRÁZEK IV.2-12: ČASOVÁ POSLOUPNOST RESIDUÍ MODELU SS MOTORKU	116

OBRÁZEK IV.2-13: GRAF RESIDUÍ V ZÁVISLOSTI NA BUDÍCÍM SIGNÁLU	116
OBRÁZEK IV.2-14: AUTOKOVARIANČNÍ FUNKCE RESIDUÍ	117
OBRÁZEK IV.3-1: SS MOTOREK S PRUŽNOU SPOJKOU - SCHÉMA SYSTÉMU	118
OBRÁZEK IV.3-2: SS MOTOR S PRUŽNOU SPOJKOU - PRINCIPIÁLNÍ SCHÉMA.....	119
OBRÁZEK IV.3-3: COULOMBOVO A VISKÓZNÍ TŘENÍ.....	119
OBRÁZEK IV.3-4: STATICKÁ CHARAKTERISTIKA SS MOTORKU S PRUŽNOU SPOJKOU	122
OBRÁZEK IV.3-5: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU SS MOTORKU S PRUŽNOU SPOJKOU NA DATECH POUŽITÝCH PRO OPTIMALIZACI	124
OBRÁZEK IV.3-6: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU SS MOTORKU S PRUŽNOU SPOJKOU NA JINÝCH DATECH NEŽ BYLA POUŽITA PRO OPTIMALIZACI	124
OBRÁZEK IV.3-7: ČASOVÁ POSLOUPNOST RESIDUÍ MODELU SS MOTORKU S PRUŽNOU SPOJKOU	125
OBRÁZEK IV.3-8: GRAF RESIDUÍ V ZÁVISLOSTI NA BUDÍCÍM SIGNÁLU	125
OBRÁZEK IV.3-9: AUTOKOVARIANČNÍ FUNKCE RESIDUÍ	126
OBRÁZEK IV.3-10: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU SS MOTORKU S PRUŽNOU SPOJKOU NA DATECH POUŽITÝCH PRO OPTIMALIZACI ...	128
OBRÁZEK IV.3-11: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU SS MOTORKU S PRUŽNOU SPOJKOU NA JINÝCH DATECH NEŽ BYLA POUŽITA PRO OPTIMALIZACI	129
OBRÁZEK IV.3-12: ČASOVÁ POSLOUPNOST RESIDUÍ MODELU SS MOTORKU S PRUŽNOU SPOJKOU	130
OBRÁZEK IV.3-13: GRAF RESIDUÍ V ZÁVISLOSTI NA BUDÍCÍM SIGNÁLU	130
OBRÁZEK IV.3-14: AUTOKOVARIANČNÍ FUNKCE RESIDUÍ SS MOTORKU S PRUŽNOU SPOJKOU	131
OBRÁZEK IV.4-1: VÝŠKA HLADINY - SCHÉMA SYSTÉMU	132
OBRÁZEK IV.4-2: PRINCIPIÁLNÍ SCHÉMA MODELU VÝŠKY HLADINY	133
OBRÁZEK IV.4-3: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU VÝŠKY HLADINY NA DATECH POUŽITÝCH PRO OPTIMALIZACI	136
OBRÁZEK IV.4-4: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU VÝŠKY HLADINY NA JINÝCH DATECH NEŽ BYLA POUŽITA PRO OPTIMALIZACI	136
OBRÁZEK IV.4-5: ČASOVÁ POSLOUPNOST RESIDUÍ MODELU VÝŠKY HLADINY	137
OBRÁZEK IV.4-6: GRAF RESIDUÍ V ZÁVISLOSTI NA BUDÍCÍM SIGNÁLU	137
OBRÁZEK IV.4-7: AUTOKOVARIANČNÍ FUNKCE RESIDUÍ SYSTÉMU VÝŠKY HLADINY...	138
OBRÁZEK IV.4-8: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU VÝŠKY HLADINY NA DATECH POUŽITÝCH PRO OPTIMALIZACI	139
OBRÁZEK IV.4-9: POROVNÁNÍ VÝSTUPU MODELU A SYSTÉMU VÝŠKY HLADINY NA JINÝCH DATECH NEŽ BYLA POUŽITA PRO OPTIMALIZACI	139
OBRÁZEK IV.4-10: ČASOVÁ POSLOUPNOST RESIDUÍ MODELU VÝŠKY HLADINY	140
OBRÁZEK IV.4-11: GRAF RESIDUÍ V ZÁVISLOSTI NA BUDÍCÍM SIGNÁLU	140
OBRÁZEK IV.4-12: AUTOKOVARIANČNÍ FUNKCE RESIDUÍ SYSTÉMU VÝŠKY HLADINY.	141

SEZNAM TABULEK

TABULKA II.2-1: POROVNÁNÍ "WHITE" A "BLACK BOX" MODELŮ	21
TABULKA II.2-2: POROVNÁNÍ MODELŮ VE SPOJITÉ A DISKRÉTNÍ OBLASTI	25
TABULKA II.2-3: STANDARDNÍ "BLACK BOX" MODELŮ.....	26
TABULKA III.9-1: VLIV PORUCH NA VÝSTUPU NA DERIVAČNÍ A INTEGRAČNÍ METODU PRO SOUSTAVU 1. ŘÁDU	87
TABULKA III.9-2: VLIV PORUCH NA VÝSTUPU NA DERIVAČNÍ A INTEGRAČNÍ METODU PRO SOUSTAVU 2. ŘÁDU	88
TABULKA III.9-3: VLIV PORUCH NA VSTUPU NA DERIVAČNÍ A INTEGRAČNÍ METODU PRO SOUSTAVU 1. ŘÁDU	89
TABULKA III.9-4: VLIV PORUCH NA VSTUPU NA DERIVAČNÍ A INTEGRAČNÍ METODU PRO SOUSTAVU 2. ŘÁDU	90
TABULKA IV.1-1: HODNOTY ÚČELOVÉ FUNKCE PRO SYSTÉM VÝMĚNÍKU.....	98
TABULKA IV.2-1: HODNOTY ÚČELOVÉ FUNKCE PRO SYSTÉM SS MOTORKU	112
TABULKA IV.3-1: HODNOTY ÚČELOVÉ FUNKCE PRO SYSTÉM SS MOTORKU S PRUŽNOU SPOJKOU	123
TABULKA IV.4-1: HODNOTY ÚČELOVÉ FUNKCE PRO SYSTÉM VÝŠKY HLADINY.....	135

SEZNAM ZÁKLADNÍCH SYMBOLŮ

a_x, b_x	koeficienty přenosu
A, B, C, D	matice stavového popisu
$A(s), B(s)$	polynomy přenosu
$\arg \min J(\mathbf{p})$	hodnota $\mathbf{p}$, která minimalizuje J
$E\{\mathbf{p}\}$	střední hodnota vektoru $\mathbf{p}$
$e(t)$	residuální odchylka
$\Phi(t)$	modulační funkce
I	jednotková matice
J	kvadratická účelová funkce
$n(t)$	šum
$n_m(t)$	měřená porucha
n_a, n_b	stupeň polynomů A, B
$\mathbf{p}$	vektor hledaných parametrů
$\mathbf{p}_{nl}$	vektor nelineárních parametrů
$\mathbf{p}_k$	vektor $\mathbf{p}$ v kroku k
q	úroveň šumu
q^{-1}	operátor zpoždění
R	regresní matice
s	Laplaceův operátor
t	nezávislá proměnná, často čas
T	vzorkovací perioda
$u(t)$	vstupní veličina
w	váhový koeficient
$y(t)$	výstupní veličina
$y_m(t)$	výstup modelu
$Y(s)$	Laplaceův obraz funkce $y(t)$
$\langle \mathbf{x}, \mathbf{y} \rangle$	skalární součin vektorů $\mathbf{x}$ a $\mathbf{y}$
$\dot{y}, y', y^{(1)}$	$\frac{dy}{dt}$
$\int_2 y(t) dt$	$\int_0^\tau \left[\int_0^\tau y(t) dt \right] dt$
z	operátor z-transformace

ABSTRACT

This dissertation addresses one of the most challenging problems in system identification, the model estimation in continuous time domain. The material is presented in three main parts. The first one reviews some basic modeling and identification concepts, algorithms and ideas from which the following parts take off. An overview of different possible model structures is given in a common framework.

The second part is focused on new identification methods in continuous time domain. In this work we prefer that the model structure is chosen by semi-physical modeling. Efficient identification algorithms and procedures for these structures are developed and analyzed. These model structures are divided into three groups:

- models linear in parameters
- nonlinear models with differentiable cost function
- nonlinear models with non-differentiable cost function

The two methods (integrational and derivational) are used for the first group of models. These methods utilize least square estimator. It is suggested to extend estimation procedure by some iterative method in order to find better value of a cost function. Further, kvazilinearization as method for the second group of models, is investigated. A combination of kvazilinearization and integrational methods was tested for partially linear in parameter model structures. This makes it possible to significantly reduce the dimension of the search space in the most application of mechatronics systems and consequently speed up search for optimal parameters. General optimization methods are used for the third group of models.

The usefulness of the proposed methods is demonstrated through a number of concrete simulation and application studies. The robustness against different kinds of perturbations is also examined.

Finally and quite importantly, the third part concerns real identification applications and analyzes usefulness of proposed methods. With these methods, high quality linear and non-linear models are estimated in continuous time domain.

I Úvodní poznámka

Výzkumná činnost v oblasti automatického řízení se v několika posledních desetiletích zaměřila především na metody identifikace v diskrétní časové oblasti. Metody pracující s modely ve spojitě časové oblasti se tak dostaly do stínu pozornosti. Přesto byly vyvinuty metody umožňující identifikaci ve spojitě časové oblasti, ale jejich použití je často spojeno se značnými omezeními (viz. metoda prof. Strejce s omezením na přechodové charakteristiky). S masivním rozvojem digitálních počítačů (v posledních 3 až 5 desetiletích) se zdá být logické, že byly vyvinuty metody hledání modelů pracující s daty v diskrétní formě (tedy stejně jak jsou prezentována data v digitálních počítačích). Tato podobnost umožnila vyvinout numericky velice efektivní algoritmy. Současně ale se zvyšováním výkonnosti komerčně využívaných počítačů můžeme dosáhnout vzorkovacích period tak malých (vzorkovací perioda několikanásobně menší než nejmenší časová konstanta identifikované technologie), že diskrétně nasnímaná data mohou nahradit spojitý signál. A společně s metodami identifikace ve spojitě časové oblasti můžeme těžit z výhod spojitých modelů v porovnání s diskrétními: u diskrétních modelů jsme omezeni vzorkovací periodou, v jejichž násobcích je počítán výstup modelu. S tím souvisí, že diskrétní systémy v sobě mohou skrývat oscilace korespondujícího spojitého systému. A konečně parametry diskrétních systémů nemívají konkrétní význam a závisí na vzorkovací periodě. Nutno ještě podotknout, že vlastnosti spojitých

a diskrétních systémů nejsou vždy analogické. V diskrétní oblasti můžeme např. realizovat oscilující soustavu prvního řádu, zatímco ve spojitě oblasti to možné není. Z toho také vyplývá obtížnost plnohodnotné transformace diskrétního popisu na spojitý. Tyto důvody nás vedly k tomu, abychom se v této práci zabývali metodami hledání modelů ve spojitě časové oblasti.

I.1 CÍLE DISERTAČNÍ PRÁCE

Cílem této práce je vyvinout a vyzkoušet metody identifikace deterministických lineárních a nelineárních modelů ve spojitě časové oblasti. V první části bude stručně shrnut postup experimentální identifikace společný identifikaci jak v diskrétní tak i ve spojitě časové oblasti a popsán současný stav poznání týkající se identifikace ve spojitě časové oblasti. V dalších částech se budeme zabývat způsoby, jak základní identifikační objekty a nástroje (data, modely, optimalizace) zapadají do procesu hledání vhodného modelu. Tato úvodní část práce bude definovat nástroje pro následné zkoumání možností identifikace ve spojitě časové oblasti.

Vlastní těžiště práce bude spočívat v metodách hledání modelů ve spojitě časové oblasti. Ty se dají rozdělit do tří skupin dle charakteru modelu:

- modely lineární v parametrech
- modely nelineární, u kterých známe parciální derivace výstupu dle vektoru hledaných parametrů
- modely nelineární, u kterých neznáme parciální derivace výstupu dle vektoru hledaných parametrů

V práci bude také stručně analyzován vliv poruch na kvalitu nalezeného modelu. Závěrečná část bude věnována praktickým ukázkám a jednotlivé metody budou vyzkoušeny na několika skutečných systémech a nalezené modely budou podrobeny kvalitativní analýze.

II Experimentální identifikace

V této kapitole jsou prezentovány základní identifikační nástroje, které jsou následně využívány v dalších částech této práce. V úvodní části (kapitola II.1) jsou definovány základní pojmy a je zde popsán všeobecný postup experimentální identifikace, tzv. identifikační cyklus. Struktury modelů, jejich vlastnosti, výhody a nevýhody jsou charakterizovány v následující kapitole (II.2) společně s porovnáním různých přístupů a oblastí používaných při modelování. Různými aspekty optimalizace účelové funkce se zabývá kapitola II.3 (od definice kritéria jakosti modelu, přes metodu nejmenších čtverců, až ke globálním metodám optimalizace). Konečný a neopomenutelný článek při modelování procesů, vyhodnocení kvality a použitelnosti nalezeného modelu je prezentován na závěr (kapitola II.4).

Naše doporučení je, že zkušenější čtenář může s klidem přeskočit tuto kapitolu a pokračovat až na straně 51 od kapitoly III, která je vlastním přínosem a těžištěm předkládané práce.

II.1 ÚVOD

II.1.1 Účel Modelování

Termíny model a modelování můžeme interpretovat mnoha různými způsoby, které se liší v závislosti od použití takového modelu. V tomto textu budeme pod pojmem „*model*“ rozumět matematický popis reálného procesu, který je vytvářený pro určité účely. Tyto účely mohou mimo jiné být:

- hluboké porozumění modelovanému systému (matematicko-fyzikální analýza),
- odhadnutí signálů, které nemůžeme měřit, z nepřímého měření,
- testování hypotéz (diagnózy v medicíně),
- předpovídání krátkodobého (adaptivní řízení) a dlouhodobého (ekonomické modely) dynamického chování,
- řízení procesů (regulace okolo nominálního bodu).

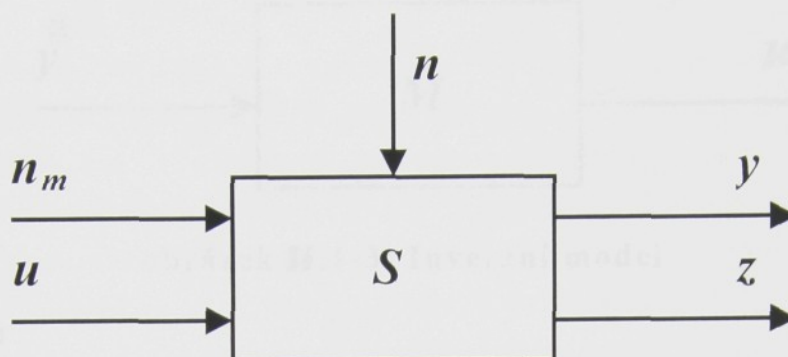
Ať je již tento účel jakýkoliv, vždy by měl být znám před začátkem modelování, protože značně ovlivňuje proceduru hledání modelu [WAL].

II.1.2 Systém

Systém (proces resp. soustava) je součástí našeho okolí, kterou jsme z něho vybrali. Považujeme ho za oddělený celek, který je ve vzájemné interakci s okolním prostředím (obrázek II.1-1):

- pozorujeme jeho vnější charakteristiky (chování) a výsledek tohoto pozorování označme jako výstup ze systému, vektor y (obrázek II.1-1),
- ze systému mohou vystupovat další charakteristiky (označme je z), které v sobě mohou zahrnovat neměřitelné signály a tím se lišit od y ,
- působíme na systém veličinami, jejichž průběh v čase známe a můžeme ho více či méně ovlivňovat. Tyto veličiny jsou tzv. vstupy do systému a označme je jako u ,

- na systém působí dva druhy poruch: ty, které můžeme měřit přímo na vstupu, nazveme jako měřené poruchy a označíme jako $\mathbf{n}_m$; ty, jejichž působení můžeme pozorovat pouze na výstupu ze systému, nazveme šumem a označíme jako $\mathbf{n}$.

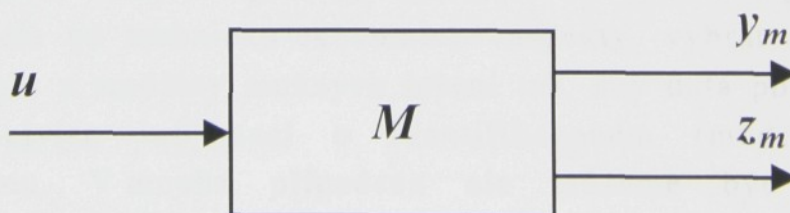


obrázek II.1-1: Systém

Pro účely modelování je rozdíl mezi vstupním signálem $\mathbf{u}$ a měřenou poruchou $\mathbf{n}_m$ nepodstatný, takže se jím již dále nebudeme zabývat a bude pro nás jen dalším vstupem do systému.

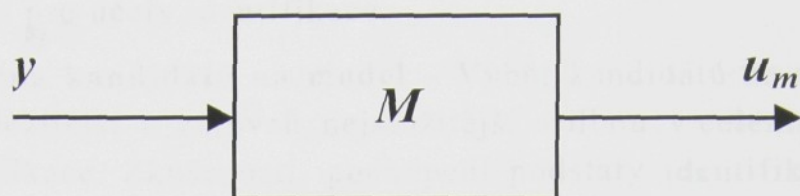
II.1.3 Model

Modelem dynamického systému rozumíme pravidlo, podle kterého můžeme z předem známých nebo změřených veličin na systému vypočítat další veličiny, které nás zajímají a které, jak doufáme, se budou dostatečně blížit skutečným hodnotám uvažovaného systému. Model nejčastěji vypočítává ze známého vstupu $\mathbf{u}$ výstup $\mathbf{y}_m$ (výstup modelu) a je navržen tak, aby se výstup modelu blížil skutečnému výstupu ze systému $\mathbf{y}$ jen jak je to možné. Zkráceně lze říci, že úkolem modelu je co nejvěrohodněji predikovat či reprodukovat chování systému. Jestliže se $\mathbf{z}$ liší od $\mathbf{y}$, může z modelu také vystupovat i vektor $\mathbf{z}_m$ (obrázek II.1-2).



obrázek II.1-2: Model

Méně tradičním přístupem, který je ale vhodné také zmínit, je model, který ze známého výstupu ze systému spočítá vektor u_m , který se co nejvíce blíží vstupu skutečnému u . Tato skupina modelů se nazývá „Inverzní modely“ (obrázek II.1-3).



obrázek II.1-3: Inverzní model

II.1.4 Fikce pravdivého modelu

Skutečný systém je jiným objektem než námi vytvořený matematický model. V určitém slova smyslu zde existuje neproniknutelná, ale průhledná přepážka mezi našim světem matematických popisů a světem skutečným. Můžeme se přes toto okno dívat a porovnávat určité vlastnosti fyzických systémů s našim matematickým popisem, nikdy ale nemůžeme tyto dva systémy zcela propojit, tak jako nemůžeme sestavit „pravdivý model“ nějakého procesu. Přijatelnost našich modelů se tak raději má řídit jejich „účelností“ než jejich „pravdivostí“ [LJU2].

II.1.5 Postup identifikace

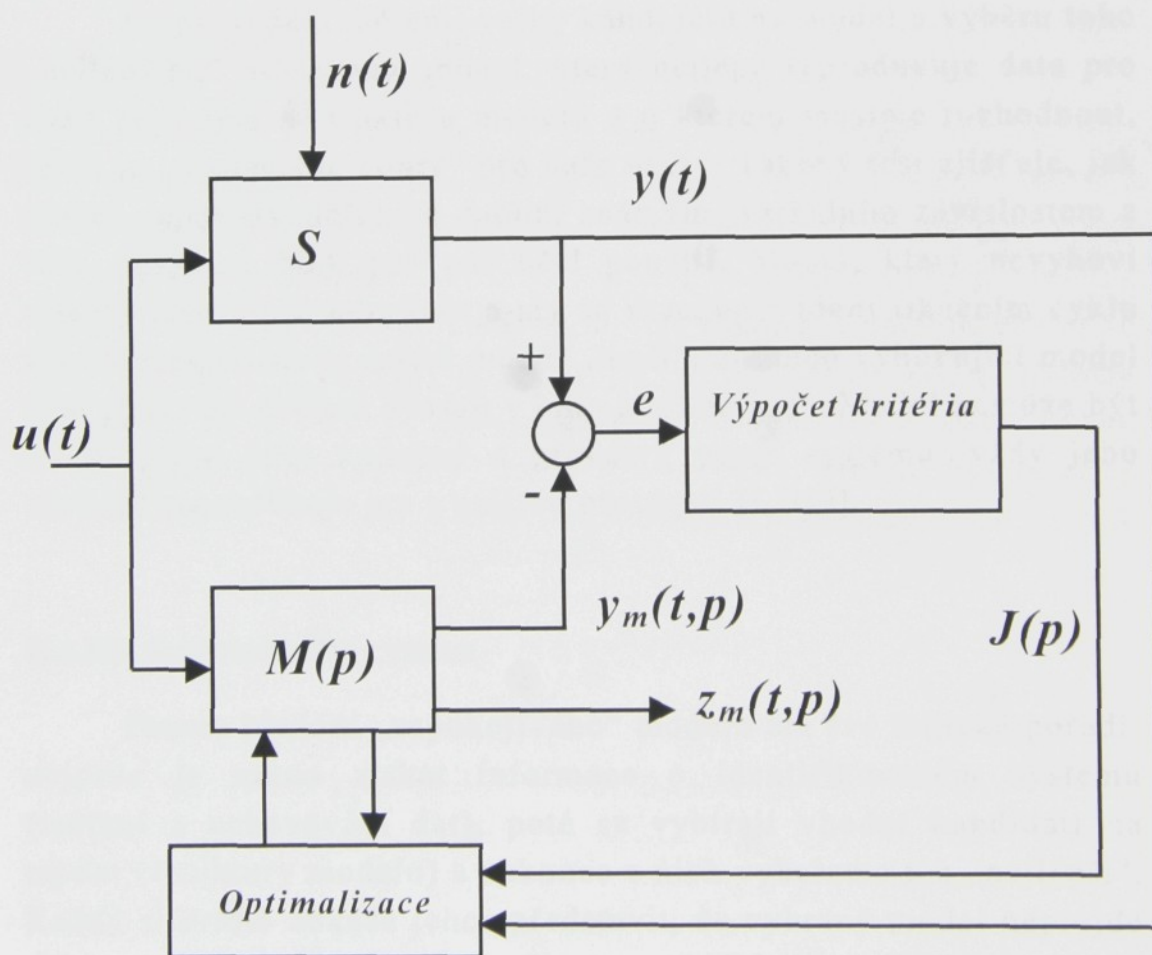
II.1.5.1 DATA, MODEL, KRITÉRIA

Konstrukce modelu z dat pracuje s třemi základními objekty:

1. **Data** – Vstupně-výstupní data jsou často nasnímána v průběhu identifikačního experimentu, při kterém má uživatel možnost rozhodnout o tom, jaký signál bude měřit a popř. také může generovat vstupní signál do systému. Úkolem uživatele je, s ohledem na technicko-ekonomické aspekty, vybrat optimální variantu z množiny možných řešení tak, aby data poskytovala co nejvíce informací o identifikovaném (modelovaném) systému. V mnoha případech ale můžeme být omezeni fyzikálními, či technologickými možnostmi systému, popř. nemáme vůbec možnost ovlivnit chování systému a musíme se spokojit jen s měřením na systému v průběhu jeho normální

činnosti modelovaného systému. To pak samozřejmě ovlivňuje naši schopnost přiblížit se ke skutečnému chování takového systému a to zejména v případech, kdy se značně liší provozní podmínky systému od podmínek, za kterých bylo prováděno měření pro účely identifikace.

2. **Množina kandidátů na model** – Výběr kandidátů na model je nejdůležitější a zároveň nejsložitější volbou v celém procesu identifikace. Zkušenosti, pochopení podstaty identifikovaného systému a v neposlední řadě i inženýrská intuice se zde musí zkombinovat pro dosažení co nejlepšího a nejjednoduššího modelu. Někdy je množina modelů vybrána až po matematicko-fyzikální analýze, kde je model s neznámými parametry sestaven ze základních fyzikálních zákonů a dalších známých závislostí (tzv. „gray box“). V jiných případech, kdy nechceme popř. ani z časových či jiných důvodů nemůžeme brát v úvahu fyzikální podstatu modelovaného systému, mohou být použity standardní lineární modely. Tento přístup k modelování se pak nazývá tzv. „black box“.



obrázek II.1-4: Tok informací při optimalizaci

3. Výběr „nejlepšího“ modelu – je vlastní identifikací. Ohodnocení kvality modelu a výběr toho „nejlepšího“ se nejčastěji provádí na základě toho jak je model schopen reprodukovat měřená data a to nejlépe data z jiného měření než data, na kterých byl model identifikován. Jak se provádí výběr nejlepšího modelu je schématicky znázorněno na následujícím obrázku: optimalizační algoritmus vypočítává ze známých informací (struktura modelu, rozdíl chování skutečného systému a modelu, trend kritéria $J(\mathbf{p}), \dots$) „nejlepší“ hodnoty optimalizovaných koeficientů $\mathbf{p}$ a to tak, aby bylo minimalizováno (popř. maximalizováno) kritérium $J(\mathbf{p})$. Tok informací při optimalizaci je zobrazen na schématu (obrázek II.1-4).

II.1.5.2 OVĚŘENÍ PLATNOSTI MODELU

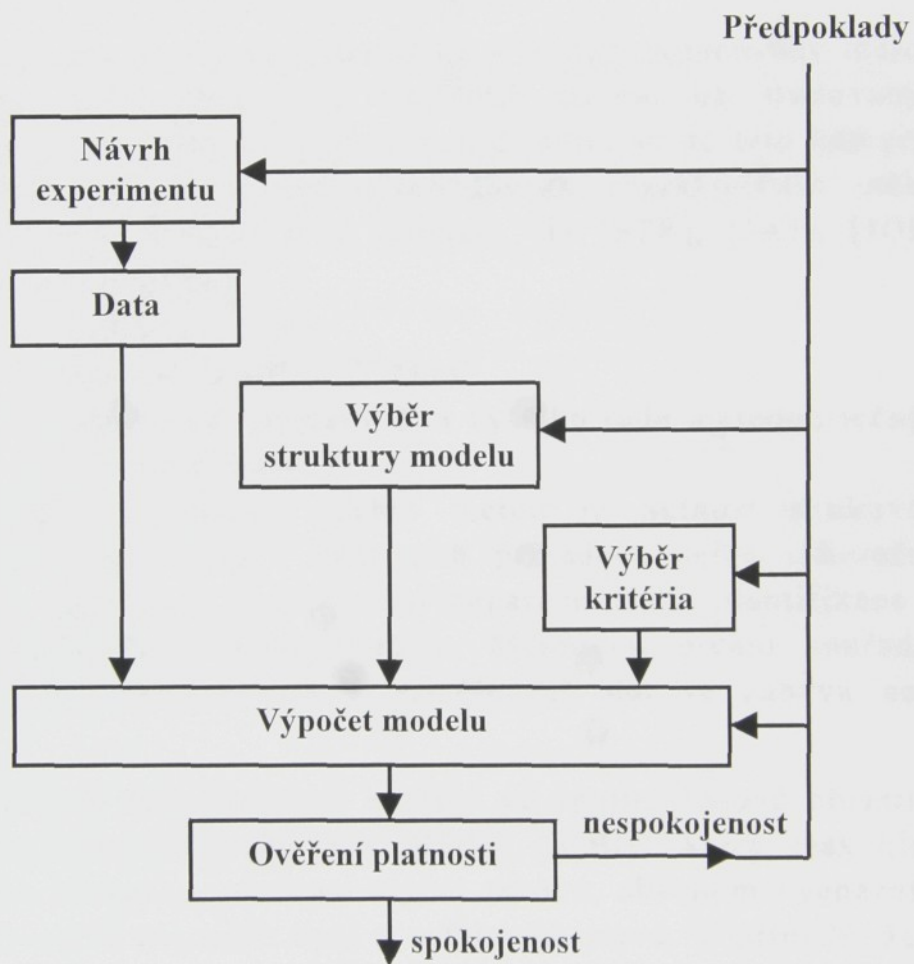
Po provedení měření, volby kandidátů na model a výběru toho „nejlepšího“ dostáváme model, který nejlépe reprodukuje data pro dané kritérium a strukturu modelu a o kterém musíme rozhodnout, jestli je „dostatečně dobrý“ pro naše účely. Takový test zjišťuje, jak model odpovídá měřeným datům, známým fyzikálním závislostem a také, jestli se hodí pro náš účel použití. Model, který nevyhoví tomuto testu bývá odmítnut a my se vracíme v identifikačním cyklu zpět a pokoušíme se nalézt model „lepší“, zatímco vyhovující model nám poskytne určitou důvěru v jeho použitelnost. Model nemůže být nikdy přijat jako konečný a pravdivý popis systému, vždy jeho chování reprodukuje jen s určitou přesností [LJU9].

II.1.5.3 IDENTIFIKAČNÍ CYKLUS

Proces hledání „uspokojivého“ modelu má své logické pořadí: nejprve je nutno získat informace o identifikovaném systému (měření a uchovávání dat), poté se vybírají vhodní kandidáti na model (struktury modelu) a nakonec z nich vybereme ten „nejlepší“. Každý si určitě dokáže lehce představit, že vybraný model neprojde testem ověření kvality, potom se musíme v identifikačním cyklu vrátit zpět a změnit některé jeho kroky (obrázek II.1-5).

Model nemusí vyhovovat z různých důvodů, uveďme např.:

- ❑ selhání numerického algoritmu hledajícího nejlepší model dle daného kritéria,
- ❑ špatně zvolené kritérium,
- ❑ nevhodná struktura modelu,
- ❑ změřená data nebyla dostatečně informativní.



obrázek II.1-5: Identifikační cyklus

II.1.6 Identifikace ve spojité časové oblasti - současný stav

Za historicky nejstarší metodu identifikace neboli určování dynamických vlastností objektů řízení ve spojité časové oblasti lze považovat matematicko-fyzikální analýzu (na základě známých údajů, fyzikálních a chemických zákonů, analytických a intuitivních aproximací sestavujeme model [SOU]). V této práci se ale zaměřujeme na identifikaci experimentální (z provozního měření vstupů a výstupů analyzovaného objektu získáme model) parametrizovaných struktur a proto se zaměříme na tuto oblast parametrické experimentální identifikace.

Jako klasické metody identifikace bývají označovány metody určené pro ruční zpracování výsledků měření na studovaných objektech. Nejdůležitější skupinu metod řadící se do této kategorie tvoří metoda vyhodnocování přechodových charakteristik (někdy označovaná jako metoda prof. Strejce, viz [STR], [JAJ], [JOH]) rozdělená na tyto případy:

- soustavy 1. řádu
- kmitavé soustavy 2. řádu
- nekmitavé soustavy 2. a vyššího řádu metodou tečny v inflexním bodě

Značnou nevýhodou těchto metod je nutnost skokového signálu na vstupu, což v některých případech nelze uskutečnit. Částečným řešením je např. použití neparametrické identifikace ke zjištění přechodové charakteristiky. Metodami určení souřadnic přechodových charakteristik z naměřených dat se zabývá např. [OLE].

Další metodou hledající modely ve spojité časové oblasti je metoda postupné integrace (viz. [JOH], [SOU]), která však klade omezení na vstupní a výstupní signál takové, abychom vycházeli z ustáleného stavu a končili opět v ustáleném stavu. V případě, že se jedná o vyhodnocení přechodových charakteristik soustav, které vznikají sériovým řazením jednodukapacitních členů, je možno použít metodu aproximace pomocí geometrického místa dominantních přenosů (viz. [SOU]). Všechny metody popsané v tomto odstavci předpokládají, že šum na výstupu soustavy je zanedbatelný.

Z tohoto výčtu je zřejmé, že hledání modelů ve spojité časové oblasti ještě stále není plně probádanou oblastí a je zde prostor pro další práci.

II.2 STRUKTURY MODELŮ

Volba struktury modelu je jedním z nejsložitějších úkolů při identifikaci systémů a intuice zde hraje velmi podstatnou roli pro návrh jednoduchého a vyhovujícího modelu. Jak již bylo zmíněno v předchozí kapitole (II.1.4), nikdy nemůžeme vytvořit tzv. „pravdivý model“ a ve velké míře záleží na nás do jaké hloubky se budeme zabývat fyzikální podstatou modelovaného systému a co se rozhodneme zanedbat.

V této kapitole budeme rozebírat různé skupiny modelů a výběr mezi těmito skupinami by se měl řídit dle:

- účelů modelování,
- pracovních podmínek, při kterých bude model použit (rozsahy a vlastnosti vstupů, šumů, ...),
- přípustných nákladů na model,
- různých dalších užitečných informací (např. informace o řídicím systému, s kterým bude model spolupracovat).

II.2.1 Fyzikální versus „black box“ modely

Fyzikální modely („white box“) se získávají tzv. matematicko-fyzikální analýzou, při které je systém rozebrán na malé součásti, pro které jsou napsány známé základní rovnice popisující jejich chování na základě fyzikálních zákonů a pravidel. Metody vytvářející takové modely vycházejí z principů zachování energie a rovnováhy sil a momentů (např. metody „Bond’s“ grafů, Lagrangeovy rovnice, ...). Tento přístup k modelování sebou přináší specifické problémy. Pro složitější systémy se počet parametrů začíná neúnosně zvyšovat a i simulace potom může být pomalá, v reálném čase i nemožná. Např. pro použití modelů některých složitých systémů pro školení obsluhy (na simulátorech), je někdy třeba je zjednodušit tak, aby mohly na současných počítačích běžet v reálném čase. Dalším specifickým problémem je to, že asi nikdy nepostihneme složitost a transparentnost přírody zcela, ale musíme si zvolit úroveň, do které se chceme ponořit (např. když si vezmeme elektrický odpor jako součást elektrického obvodu, tak ho můžeme modelovat jednoduchým vztahem (Ohmovým zákonem), jestliže ale

půjdeme o úroveň níže, zjistíme, že na skutečný odpor se můžeme dívat jako na odpor a cívku zapojené v sérii, ...).

„**Black box**“ modely stojí na opačné straně spektra. Tyto modely aproximují chování systému bez jakýkoliv nároků na informace o systému vlastním. Není ani důležité, co jednotlivé vstupy a výstupy ve skutečnosti znamenají a v jakých jednotkách se udávají. Parametry těchto modelů nemají žádný fyzikální význam. Např. jakékoliv měření můžeme s určitou přesností aproximovat polynomem $y_m = p_1 + p_2t + p_3t^2 + \dots$, jestliže stupeň polynomu bude dostatečně vysoký. Takový model bude samozřejmě mít velice špatné schopnosti předvídat chování obecného dynamického systému, protože vůbec nerespektuje vstupy do něho. Samozřejmě ale existují i vyspělejší struktury „black box“ modelů (viz. např. [SOD], [LJU2]), které mohou uspokojivě pokrýt skupiny „lineárních“ dynamických systémů.

Vlastnosti těchto dvou skupin modelů jsou shrnuty v následující tabulce (tabulka II.2-1).

MODELÝ		
	"White box"	"Black box"
Parametry	mají konkrétní význam	nemají konkrétní význam
Simulace	dlouhá a složitá	rychlá a jednoduchá
Fyzikální informace o systému	mohou být zahrnuty viditelně	většinou nebrány v úvahu
Oblast platnosti modelu	velká (jestliže je správná struktura)	omezená

tabulka II.2-1: Porovnání "white" a "black box" modelů

II.2.2 Lineární versus nelineární modely

Rozlišme dva typy linearit:

- modely lineární ve vstupech
- modely lineární v parametrech

Model je **lineární ve vstupech**, jestliže jeho výstupy splňují princip superpozice s ohledem na jeho vstupy, což znamená, jestliže platí:

$$\mathbf{y}_m(t, \mathbf{p}, \lambda \mathbf{u}_1 + \mu \mathbf{u}_2) = \lambda \mathbf{y}_m(t, \mathbf{p}, \mathbf{u}_1) + \mu \mathbf{y}_m(t, \mathbf{p}, \mathbf{u}_2), \quad (\text{II-1})$$

kde $\mathbf{y}_m(t, \mathbf{p}, \mathbf{u})$ je výstup modelu v čase t , s parametry $\mathbf{p}$, se vstupy $\mathbf{u}$ a s nulovými počátečními podmínkami.

Když inženýři pracující v oblasti technické kybernetiky hovoří o lineárních modelech, mají většinou na mysli tento typ linearity. Navíc ještě většinou předpokládají, že se jedná o model časově nezávislý, což znamená, že jeho parametry nejsou závislé na posunu časové osy.

Model je **lineární v parametrech**, jestliže jeho výstupy splňují princip superpozice s ohledem na jeho parametry, to znamená, jestliže platí:

$$\mathbf{y}_m(t, \lambda \mathbf{p}_1 + \mu \mathbf{p}_2, \mathbf{u}) = \lambda \mathbf{y}_m(t, \mathbf{p}_1, \mathbf{u}) + \mu \mathbf{y}_m(t, \mathbf{p}_2, \mathbf{u}). \quad (\text{II-2})$$

Pro lepší názornost se podívejme na několik příkladů:

- | | |
|-------------------------------------|--|
| $y_m(t+1, p) = pu(t)$ | je lineární jak v parametrech tak i ve vstupech, |
| $y_m(t+1, p) = py_m(t, p) + u(t)$ | je lineární ve vstupech, ale ne v parametrech, |
| $y_m(t+1, p) = pu^2(t)$ | je lineární v parametrech, ale ne ve vstupech, |
| $y_m(t+1, p) = py_m^2(t, p) + u(t)$ | není lineární ani vstupech ani v parametrech. |

Tam, kde to jde, se při modelování preferují lineární struktury před nelineárními. Pro modely lineární ve vstupech existuje rozsáhlá matematická teorie (stabilita, optimální řízení, robustní řízení, adaptivní řízení, ...) a proto je tento popis zvláště výhodný pro oblast technické kybernetiky. Nalezení parametrů modelů lineárních v parametrech je většinou jednoduché a mnohdy je možné použít explicitního vztahu a tím se vyhnout iterování při hledání „nejlepšího“ modelu z množiny kandidátů [FOR1].

Na druhou stranu se v praxi těžko setkáme se systémem, který by byl lineární v celém rozsahu vstupně-výstupních veličin a proto mají lineární struktury jen omezenou oblast použitelnosti. Proto nemůžeme nelineární struktury úplně vyloučit z okruhu našeho zájmu. Osvědčí se např. tam, kde řízení navrhujeme na linearizovaném modelu, ale simulaci chování provádíme na modelu nelineárním.

II.2.3 Modely ve spojitě versus v diskrétní oblasti

Modelované soustavy převážně existují ve spojitě časové oblasti. Proto také tradiční přístup k modelování je popis systémů diferenciálními rovnicemi, zvláště pak stavovými modely:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}(\mathbf{p})\mathbf{x}(t) + \mathbf{B}(\mathbf{p})\mathbf{u}(t), \quad \mathbf{x}(0) = \mathbf{x}_0(\mathbf{p}), \\ \mathbf{y}_m(t) &= \mathbf{C}(\mathbf{p})\mathbf{x}(t) + \mathbf{D}(\mathbf{p})\mathbf{u}(t).\end{aligned}\tag{II-3}$$

Z tohoto typu popisu (II-3) pak můžeme snadno přejít na popis pomocí přenosových matic:

$$\begin{aligned}\mathbf{y}_m(s, \mathbf{p}) &= \mathbf{H}_1(s, \mathbf{p})\mathbf{u}(s) + \mathbf{H}_2(s, \mathbf{p})\mathbf{x}_0(\mathbf{p}), \\ \mathbf{H}_1(s, \mathbf{p}) &= \mathbf{C}(\mathbf{p})[s\mathbf{I} - \mathbf{A}(\mathbf{p})]^{-1}\mathbf{B}(\mathbf{p}) + \mathbf{D}(\mathbf{p}), \\ \mathbf{H}_2(s, \mathbf{p}) &= \mathbf{C}(\mathbf{p})[s\mathbf{I} - \mathbf{A}(\mathbf{p})]^{-1},\end{aligned}\tag{II-4}$$

kde s je Laplaceův operátor a $\mathbf{I}$ jednotková matice. A konečně se můžeme vrátit do časové oblasti a přejít na popis pomocí soustavy diferenciálních rovnic:

$$\sum_{i=0}^n a_i^{(j)} y_m^{(j)} = \sum_{i=0}^m b_i^{(j)} u^{(j)}.\tag{II-5}$$

Masové používání číslicových počítačů znamenalo veliký rozmach teorie řízení v diskrétní časové oblasti. Simulace diskrétních modelů na číslicových počítačích je mnohem rychlejší a jednodušší než u spojitých modelů, což je předurčuje pro řízení v reálném čase. Avšak používání diskrétních modelů může způsobit ztrátu některých informací o chování modelovaného spojitého systému [WAL].

Podobně, jak jsme to provedli ve spojité oblasti, tak i v diskrétní oblasti zavedeme stavový popis:

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}(\mathbf{p})\mathbf{x}(t) + \mathbf{B}(\mathbf{p})\mathbf{u}(t), \quad \mathbf{x}(0) = \mathbf{x}_0(\mathbf{p}), \\ \mathbf{y}_m(t) &= \mathbf{C}(\mathbf{p})\mathbf{x}(t) + \mathbf{D}(\mathbf{p})\mathbf{u}(t). \end{aligned} \quad (\text{II-6})$$

A analogicky můžeme přejít na popis pomocí přenosových matic:

$$\begin{aligned} \mathbf{y}_m(z, \mathbf{p}) &= \mathbf{H}_1(z, \mathbf{p})\mathbf{u}(z) + \mathbf{H}_2(z, \mathbf{p})\mathbf{x}_0(\mathbf{p}), \\ \mathbf{H}_1(z, \mathbf{p}) &= \mathbf{C}(\mathbf{p})[\mathbf{z}\mathbf{I} - \mathbf{A}(\mathbf{p})]^{-1}\mathbf{B}(\mathbf{p}) + \mathbf{D}(\mathbf{p}), \\ \mathbf{H}_2(z, \mathbf{p}) &= \mathbf{C}(\mathbf{p})[\mathbf{z}\mathbf{I} - \mathbf{A}(\mathbf{p})]^{-1}\mathbf{z}. \end{aligned} \quad (\text{II-7})$$

Tento popis můžeme jednoduše přepsat ve tvaru soustavy rekurentních diferenčních rovnic:

$$\sum_{i=-1}^{n-1} a_i y_m(t-i) = \sum_{k=0}^{m-1} b_k u(t-k). \quad (\text{II-8})$$

Diferenční rovnice se často píšou ve zhuštěnější formě, kde se používá operátor zpoždění $q^{-1}x(t) = x(t-1)$:

$$A(q^{-1})y_m(t+1) = B(q^{-1})u(t+1). \quad (\text{II-9})$$

Z důvodu snadnější simulace diskrétních systémů lze posunout hranice experimentálního studování jejich vlastností podstatně dále. Na druhou stranu jsme omezeni vzorkovací periodou, v jejích násobcích je počítán výstup modelu. Navíc v sobě mohou diskrétní systémy skrývat oscilace přidruženého spojitého systému. A konečně parametry diskrétních systémů nemívají konkrétní význam (závisí na vzorkovací periodě) a vlastnosti spojitých a diskrétních systémů nejsou vždy analogické.

Vlastnosti těchto dvou skupin modelů jsou shrnuty na následujícím obrázku (tabulka II.2-2).

MODELÝ		
	<i>Spojité oblast</i>	<i>Diskrétní oblast</i>
Parametry	nezávisí na vzorkovací periodě	závisí na vzorkovací periodě
Simulace	vyžaduje diskretizaci	jednoduchá
Fyzikální informace o systému	mohou být zahrnuty viditelně	většinou nebrány v úvahu (zesílení výjimkou)
Časy vzorkování	nezávisle vybírány	dané vzorkovací periodou

tabulka II.2-2: Porovnání modelů ve spojitě a diskrétní oblasti

II.2.4 Deterministické versus stochastické modely

Modely popsané v kapitole II.2.3 patří do skupiny deterministických systémů, což znamená, že popisují výstupy ze systému pouze v závislosti na vstupních veličinách a neuvažují přitom s náhodnými poruchami. Tento popis je v praxi mnohdy nedostačující, protože na systém působí různé poruchy a i měření se provádí s určitou chybou. Proto byla definována další skupina modelů, tzv. stochastické modely, kde jsou poruchy popsány pomocí teorie pravděpodobnosti, tzv. stochastickými procesy (posloupnostmi náhodných proměnných).

Pro jednoduchost budeme předpokládat systém s jedním vstupem a jedním výstupem popsaný diferenční rovnicí :

$$y(t) = G(q, p)u(t) + H(q, p)n(t), \quad (\text{II-10})$$

kde $G(q, p)$ a $H(q, p)$ jsou přenosové funkce. V závislosti na tom, jakým způsobem se projevuje porucha na výstupu ze systému, mohou být definovány různé struktury modelů. Obecná struktura modelu, ze které následovně vycházejí další, je definována ve vztahu (II-11):

$$y(t) = q^{-n_k} \frac{B(q, \mathbf{p})}{A(q, \mathbf{p})F(q, \mathbf{p})} u(t) + \frac{C(q, \mathbf{p})}{A(q, \mathbf{p})D(q, \mathbf{p})} e(t), \quad (\text{II-11})$$

kde n_k určuje dopravní zpoždění mezi vstupním a výstupním signálem a charakteristické polynomy jsou definovány:

$$A(q, \mathbf{p}) = 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a}, \quad (\text{II-12})$$

$$B(q, \mathbf{p}) = b_1 + b_2 q^{-1} + \dots + b_{n_b} q^{-n_b}, \quad (\text{II-13})$$

$$C(q, \mathbf{p}) = 1 + c_1 q^{-1} + \dots + c_{n_c} q^{-n_c}, \quad (\text{II-14})$$

$$D(q, \mathbf{p}) = 1 + d_1 q^{-1} + \dots + d_{n_d} q^{-n_d}, \quad (\text{II-15})$$

$$F(q, \mathbf{p}) = 1 + f_1 q^{-1} + \dots + f_{n_f} q^{-n_f}. \quad (\text{II-16})$$

Vektor hledaných parametrů $\mathbf{p}$ je definován:

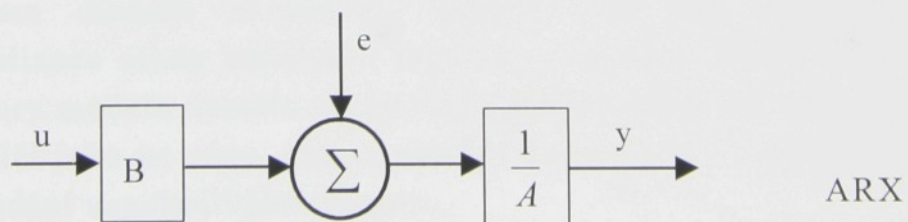
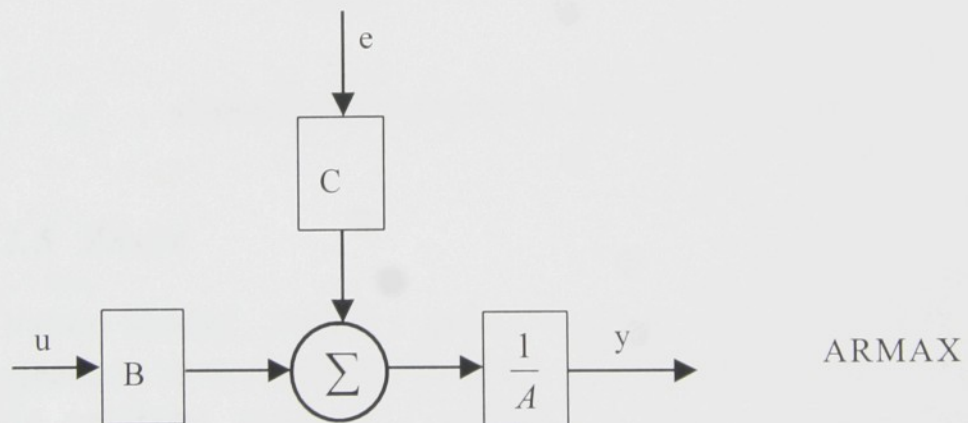
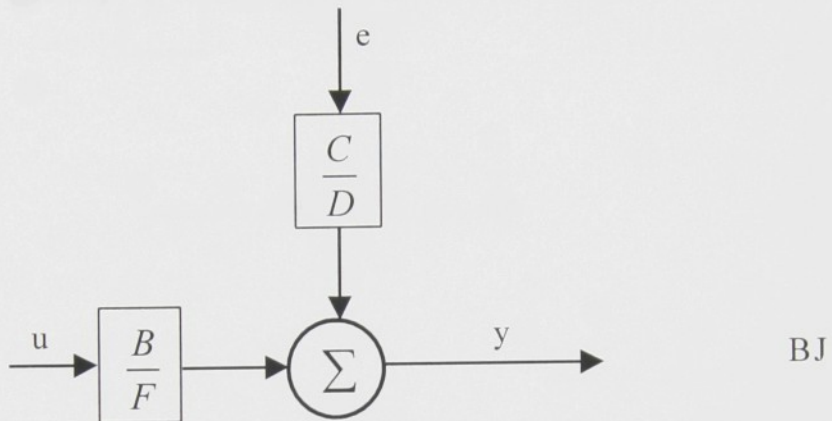
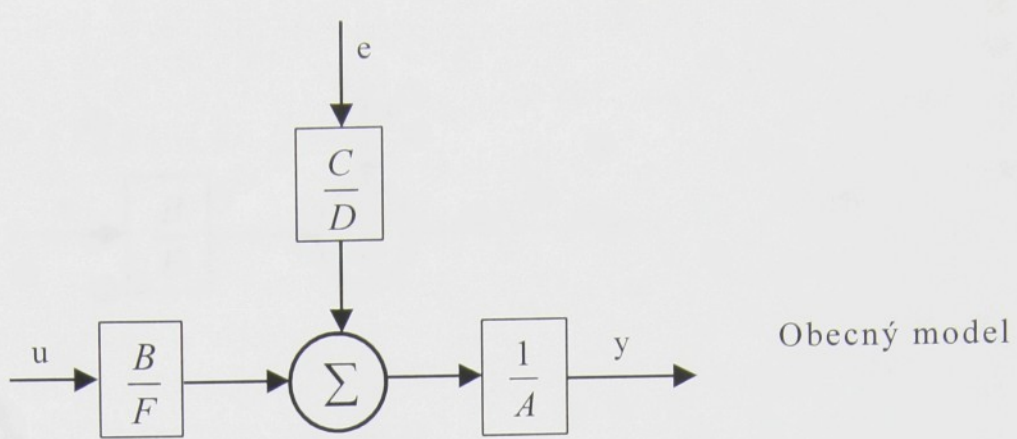
$$p = [a_1, \dots, a_{n_a}, \dots, f_1, \dots, f_{n_f}].$$

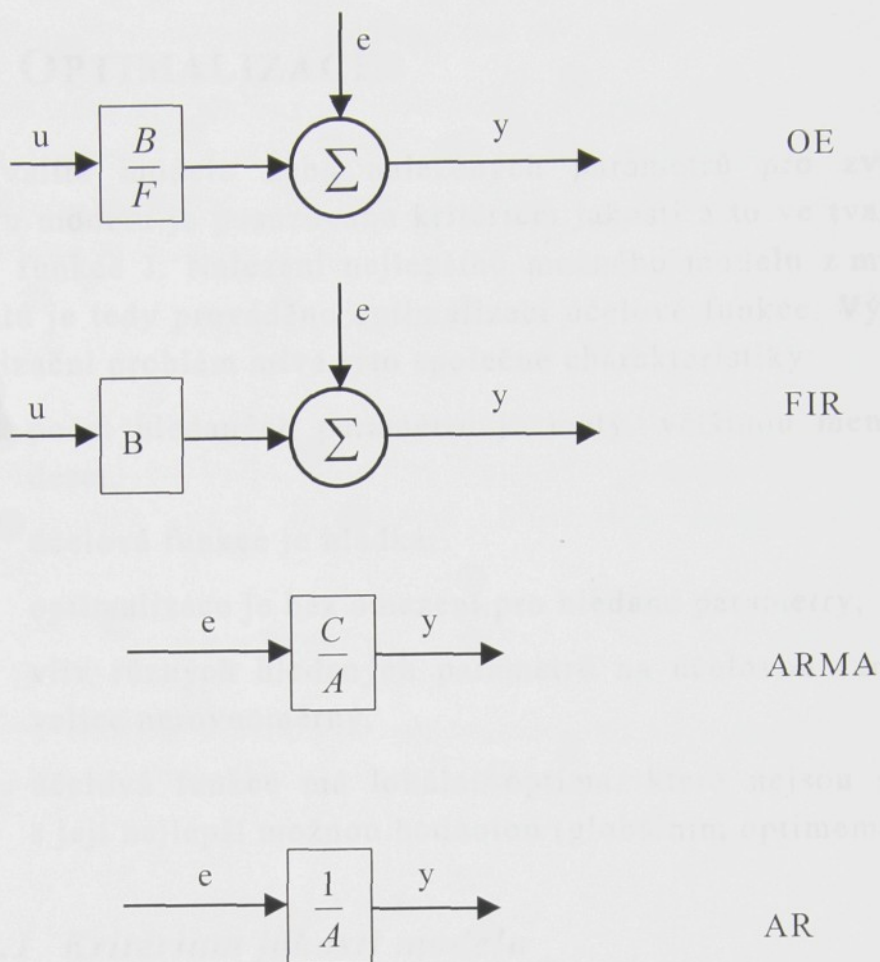
Některé struktury modelů získané zjednodušením obecné struktury dostaly pro svojí úspěšnost vlastní jména (tabulka II.2-3):

Název	Zjednodušené polynomy
FIR	$A=C=D=F=1$
ARX	$C=D=F=1$
OE	$A=C=D=1$
ARMAX	$D=F=1$
AR	$B=0, C=D=1$
ARMA	$B=0, D=1$
BJ	$A=1$

tabulka II.2-3: Standardní "black box" modely

Blokové vyjádření těchto struktur je na následujícím schématu (obrázek II.2-1).





obrázek II.2-1: Struktury modelů

II.2.5 Závěr

Výběr vhodné struktury modelu nelze zobecnit. Vždy se pohybujeme ve více či méně otevřené množině přípustných řešení, ze které vybíráme vhodnou strukturu modelu. Navíc při vlastní optimalizaci můžeme nalézt „nejlepší“ model jen z omezené množiny dané definovanou strukturou modelu. Jestliže byla struktura modelu nevhodně vybrána, pak ani nejdůmyslnější optimalizace nikdy nenalezne uspokojivý model. Proto by se volba struktury modelu neměla nikdy brát jako konečná, ale mělo by se na ni hledět jako na něco, čím je vhodné se zabývat i v dalších krocích při hledání uspokojivého modelu.

V další kapitole budeme již považovat strukturu modelu za vybranou a budeme se zabývat pouze tím, jak nalézt „nejlepší“ model mezi těmito kandidáty.

II.3 OPTIMALIZACE

Kvalita modelu nebo nalezených parametrů pro zvolenou strukturu modelu je posuzována kritériem jakosti a to ve tvaru tzv. účelové funkce J . Nalezení nejlepšího možného modelu z množiny kandidátů je tedy prováděno optimalizací účelové funkce. Výsledný optimalizační problém má tyto společné charakteristiky:

- počet hledaných parametrů je malý, většinou menší než deset,
- účelová funkce je hladká,
- optimalizace je bez omezení pro hledané parametry,
- vliv různých hledaných parametrů na účelovou funkci je velice nerovnoměrný,
- účelová funkce má lokální optima, která nejsou shodná s její nejlepší možnou hodnotou (globálním optimem).

II.3.1 Kritérium jakosti modelu

Kritériem výběru „nejlepšího“ modelu je v našem případě účelová funkce $J(\mathbf{p})$ vyhodnocovaná pro parametry modelu $\mathbf{p}$. Je zřejmé, že optimální parametry modelu $\mathbf{p}$ závisí na volbě účelové funkce $J(\mathbf{p})$.

Označme vektor $\mathbf{y}$ jako výsledek měření na identifikovaném systému a vektor $\mathbf{y}_m(\mathbf{p})$ jako výstup hledaného modelu $\mathbf{M}(\mathbf{p})$. Tyto vektory mohou být psány následovně:

$$\mathbf{y} = \begin{bmatrix} y(t_1) \\ y(t_2) \\ \dots \\ y(t_n) \end{bmatrix} \quad \text{a} \quad \mathbf{y}_m(\mathbf{p}) = \begin{bmatrix} y_m(t_1, \mathbf{p}) \\ y_m(t_n, \mathbf{p}) \\ \dots \\ y_m(t_n, \mathbf{p}) \end{bmatrix}$$

V důsledku působících poruch na systém většinou není možné dosáhnout stejného výsledku měření $\mathbf{y}$, jestliže experiment opakujeme, proto je $\mathbf{y}$ náhodný vektor.

II.3.1.1 KRITÉRIUM NEJMENŠÍCH ČTVERCŮ

Kvadratické účelové funkce jsou nejpoužívanějším kritériem vůbec a to pro relativně jednoduchou optimalizaci (pro lineární modely lze vypočítat optimum kvadratické účelové analyticky). Tyto účelové funkce lze psát:

$$J(\mathbf{p}) = \mathbf{e}^T(\mathbf{p})\mathbf{Q}\mathbf{e}(\mathbf{p}), \quad (\text{II-17})$$

kde $\mathbf{Q}$ je kladná symetrická váhová matice a $\mathbf{e}$ je vektor charakterizující odchylku chování systému a modelu. Optimalizací účelové funkce poté dostaneme optimální hodnotu vektoru parametrů:

$$\hat{\mathbf{p}} = \arg \min J(\mathbf{p}). \quad (\text{II-18})$$

Tato metoda se nazývá *vážená metoda nejmenších čtverců*. Velmi často se odchylka chování systému a modelu charakterizuje takto:

$$\mathbf{e}(\mathbf{p}) = \mathbf{y} - \mathbf{y}_m(\mathbf{p}) \quad (\text{II-19})$$

a jestliže zvolíme matici $\mathbf{Q}$ jako diagonální, můžeme účelovou funkci psát následovně:

$$J(\mathbf{p}) = \frac{1}{N} \sum_k \sum_{t_{ik}} w_{ik} [y_k(t_{ik}) - y_{m_i}(t_{ik}, \mathbf{p})]^2, \quad (\text{II-20})$$

kde N je celkový počet experimentálních dat. Váhové koeficienty w_{ik} jsou kladné, popř. nulové a odpovídají prvkům ležícím na hlavní diagonále matice $\mathbf{Q}$. Čím vyšší w_{ik} , tím vyšší zaplatíme cenu (zvýší se účelová funkce) za odchylku modelu od experimentálních dat $y_k(t_{ik})$. Proto např.:

- $w_{ik}=0$ zanedbá hodnotu považovanou za nevýznamnou
- $w_{ik}=[y_k(t_{ik})]^{-2}$ (za předpokladu $y_k(t_{ik}) \neq 0$) způsobí, že chyba se stane relativní a zlepší se přiblížení výstupů modelu a systému u malých hodnot $y_k(t_{ik})$.

Váhový koeficient může být vybírán iteračně a jeho zvětšení v oblasti, kterou považujeme za nejdůležitější, můžeme dosáhnout uspokojivého kompromisu.

II.3.1.2 KRITÉRIUM ABSOLUTNÍCH HODNOT

Přestože se metoda nejmenších čtverců stala nejpoužívanější, neměli bychom se omezovat pouze na ni. Můžeme např. chtít minimalizovat váženou sumu absolutních hodnot odchylek modelu od měření:

$$J(\mathbf{p}) = \frac{1}{N} \sum_k \sum_{t_{ik}} w_{ik} |y_k(t_{ik}) - y_{m_k}(t_{ik}, \mathbf{p})|, \quad (\text{II-21})$$

taková účelová funkce, která již byla zmíněna v práci Galilea (v roce 1632!), trestá velmi velké chyby méně než kvadratická účelová funkce.

Nevýhodou této účelové funkce je to, že neexistuje analytické řešení optimálních parametrů a to z důvodu obtížné diferencovatelnosti takto definované účelové funkce.

II.3.1.3 KRITÉRIUM MAXIMÁLNÍ VĚROHODNOSTI

Známe-li předem hustotu pravděpodobnosti náhodného vektoru $\mathbf{e}$, můžeme použít kritérium maximální věrohodnosti:

$$J(\mathbf{p}) = \pi(\mathbf{e}, \mathbf{p}). \quad (\text{II-22})$$

Při této metodě se hledá *maximum* účelové funkce, která je definována jako podmíněná hustota pravděpodobnosti. Vektor $\mathbf{y}$ jako měření na systému je dán, proto se lze dívat na hustotu pravděpodobnosti $\pi(\mathbf{e}, \mathbf{p})$ jako na funkci závisující na parametrech modelu $\mathbf{p}$. Na $\pi(\mathbf{e}, \mathbf{p})$ se lze potom dívat jako na věrohodnost vektoru $\mathbf{e}$. Metoda maximální věrohodnosti hledá takové hodnoty vektoru parametrů $\mathbf{p}$, které odpovídají maximální věrohodnosti měřených dat.

Tento přístup nám umožňuje zahrnout do identifikačního cyklu informace o poruchách působících na systém.

V praxi se pro zjednodušení výpočtu obvykle zavádí účelová funkce ve tvaru:

$$J(\mathbf{p}) = \ln \pi(\mathbf{e}, \mathbf{p}), \quad (\text{II-23})$$

která má maximum pro stejné hodnoty parametrů $\mathbf{p}$ jako původní účelová funkce.

II.3.2 Metoda nejmenších čtverců

Metoda nejmenších čtverců vyvinutá na počátku 19. století minimalizuje účelovou funkci ve tvaru:

$$J(\mathbf{p}) = \mathbf{e}^T(\mathbf{p})\mathbf{Q}\mathbf{e}(\mathbf{p}), \quad (\text{II-24})$$

kde: $\mathbf{e}(\mathbf{p}) = \mathbf{y} - \mathbf{R}\mathbf{p}.$ (II-25)

Optimum účelové funkce nalezneme položením její první derivace dle vektoru hledaných parametrů rovno nule:

$$\left. \frac{\partial J}{\partial \mathbf{p}} \right|_{\mathbf{p}=\hat{\mathbf{p}}} = -2\mathbf{R}^T \mathbf{Q}(\mathbf{y} - \mathbf{R}\hat{\mathbf{p}}) = 0. \quad (\text{II-26})$$

A za předpokladu, že existuje inverze $\mathbf{R}^T \mathbf{Q} \mathbf{R}$, lze psát pro optimální hodnoty parametrů:

$$\hat{\mathbf{p}} = (\mathbf{R}^T \mathbf{Q} \mathbf{R})^{-1} \mathbf{R}^T \mathbf{Q} \mathbf{y}. \quad (\text{II-27})$$

II.3.2.1 NUMERICKÉ ŘEŠENÍ METODY NEJM. ČTVERCŮ

Řešení vztahu (II-26) lze provést různými numerickými metodami. Hlavním problémem při řešení tohoto vztahu dle (II-27) pro odhad parametrů modelu je výpočet inverze matice $\mathbf{R}^T \mathbf{R}$ (pro jednoduchost uvažujme matici $\mathbf{Q}$ jako jednotkovou). Pro špatně podmíněné problémy, kdy je matice $\mathbf{R}^T \mathbf{R}$ blízká singulární, existují numerické algoritmy (QR, Choleského dekompozice ...) založené na dekompozici této matice. Myšlenkou těchto metod je, že matice $\mathbf{R}^T \mathbf{R}$ by se neměla vytvářet, protože obsahuje součin originálních dat.

Jednoduché pravidlo pro test matice $\mathbf{R}$ je: jestliže poměr nejmenšího a největšího singulárního čísla matice $\mathbf{R}$ je menší než přesnost plovoucí čárky počítače (asi 10^{-16}), pak je problém špatně podmíněn a výsledky získané ze vztahu (II-27) budou s vysokou pravděpodobností zatíženy velkou chybou.

QR dekompozice:

Přestože QR dekompozice existuje pro obecnou obdélníkovou matici, my se zde pro jednoduchost omezíme na případ, kdy všechny matice jsou čtvercové rozměru $N \times N$.

Hledáme řešení rovnice

$$\mathbf{b} = \mathbf{A} \cdot \mathbf{x}, \quad (\text{II-28})$$

matici $\mathbf{A}$ rozložíme tak, že

$$\mathbf{A} = \mathbf{Q} \cdot \mathbf{R}, \quad (\text{II-29})$$

kde $\mathbf{R}$ je horní trojúhelníková matice a $\mathbf{Q}$ je čtvercová matice, taková, že platí:

$$\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{I},$$

potom již stačí dosadit (II-29) do (II-28) a výsledek vynásobit maticí $\mathbf{Q}^T$ a dostaneme:

$$\mathbf{Q}^T \cdot \mathbf{b} = \mathbf{Q}^T \cdot \mathbf{Q} \cdot \mathbf{R} \cdot \mathbf{x}$$

a po úpravě:

$$\mathbf{Q}^T \cdot \mathbf{b} = \mathbf{R} \cdot \mathbf{x}$$

a protože matice $\mathbf{R}$ je horní trojúhelníková, řešení neznámého vektoru $\mathbf{x}$ dostaneme dosazením od nejspodnějšího řádku.

Metodami hledání matic $\mathbf{Q}$ a $\mathbf{R}$ se v tomto textu zabývat nebudeme, standardní softwarové balíky je již mají implementovány a i literatura dotýkající se tohoto problému je rozsáhlá (viz. např. [GUN], [LJU13], [NIU]).

II.3.3 Rekurzivní metoda nejmenších čtverců

Algoritmus metody nejmenších čtverců tak, jak byl prezentován v předchozích kapitolách, potřebuje všechna data, která byla změřena na systému najednou (off-line), zatímco její rekurzivní varianta zpracovává data z měření po sobě. Tento přístup může být výhodný hned z několika důvodů:

- k dispozici není dostatečná paměť počítače, tudíž nelze data zaznamenat a poté najednou spustit optimalizaci.

Můžeme si např. přát uskladňovat pouze omezená množství dat nezávislá na počtu měřených vzorků,

- dalším důvodem může být potřeba znát hodnoty parametrů identifikovaného systému průběžně, např. pro adaptivní regulaci či pro detekci, jestli jsou parametry systému v daném intervalu (např. pro bezpečnost výrobní linky).

II.3.3.1 VEKTOR $\mathbf{p}$ SE PŘEDPOKLÁDÁ KONSTANTNÍ

Pro jednoduchost uvažujme váhovou matici $\mathbf{Q}$ jako jednotkovou. Následně můžeme pro nerekurzivní metodu nejmenších čtverců psát vztah pro vektor optimálních parametrů ve tvaru:

$$\hat{\mathbf{p}}(t) = [\mathbf{R}^T(t-1)\mathbf{R}(t-1)]^{-1}\mathbf{R}^T(t-1)\mathbf{y}(t), \quad (\text{II-30})$$

kde:

$$\mathbf{R}(t-1) = \begin{bmatrix} \mathbf{r}^T(0) \\ \mathbf{r}^T(1) \\ \dots \\ \mathbf{r}^T(t-1) \end{bmatrix} \quad \mathbf{a} \quad \mathbf{y}(t) = \begin{bmatrix} y(0) \\ y(1) \\ \dots \\ y(t) \end{bmatrix} \quad (\text{II-31})$$

v čase $t+1$ změříme další hodnotu výstupu ze systému $y(t+1)$ a přejeme si tuto informaci zahrnout do výpočtu optimálních parametrů, ale bez toho abychom si museli pamatovat staré hodnoty $\mathbf{R}$ a $\mathbf{y}$, proto píšeme:

$$\hat{\mathbf{p}}(t) = \left[\sum_{i=1}^t \mathbf{r}(i-1)\mathbf{r}^T(i-1) \right]^{-1} \left[\sum_{i=1}^t \mathbf{r}(i-1)y(i) \right] \quad (\text{II-32})$$

a definujeme:

$$\mathbf{M}(t) = \sum_{i=1}^t \mathbf{r}(i-1)\mathbf{r}^T(i-1) \quad \mathbf{a} \quad \mathbf{v}(t) = \sum_{i=1}^t \mathbf{r}(i-1)y(i),$$

potom:

$$\hat{\mathbf{p}}(t) = \mathbf{M}^{-1}(t)\mathbf{v}(t), \quad (\text{II-33})$$

což můžeme psát v rekurzivním tvaru:

$$\mathbf{M}(t+1) = \mathbf{M}(t) + \mathbf{r}(t)\mathbf{r}^T(t), \quad (\text{II-34})$$

$$\mathbf{v}(t+1) = \mathbf{v}(t) + \mathbf{r}(t)y(t+1), \quad (\text{II-35})$$

$$\hat{\mathbf{p}}(t+1) = \mathbf{M}^{-1}(t+1)\mathbf{v}(t+1). \quad (\text{II-36})$$

Proto si dostačuje pamatovat hodnoty $\mathbf{M}$ a $\mathbf{v}$, což samozřejmě vyžaduje mnohem méně paměti než ukládání $\mathbf{R}$ a $\mathbf{y}$. Dále vyjádříme $\mathbf{p}(t+1)$ jako funkci $\mathbf{p}(t)$:

$$\hat{\mathbf{p}}(t+1) = \mathbf{M}^{-1}(t+1)[\mathbf{v}(t) + \mathbf{r}(t)y(t+1)] = \mathbf{M}^{-1}(t+1)[\mathbf{M}(t)\hat{\mathbf{p}}(t) + \mathbf{r}(t)y(t+1)],$$

$$\hat{\mathbf{p}}(t+1) = \mathbf{M}^{-1}(t+1)\{\mathbf{M}(t+1) - \mathbf{r}(t)\mathbf{r}^T(t)\}\hat{\mathbf{p}}(t) + \mathbf{r}(t)y(t+1), \quad (\text{II-37})$$

po zjednodušení dostaneme:

$$\hat{\mathbf{p}}(t+1) = \hat{\mathbf{p}}(t) + \mathbf{M}^{-1}(t+1)\mathbf{r}(t)[y(t+1) - \mathbf{r}^T(t)\hat{\mathbf{p}}(t)]. \quad (\text{II-38})$$

Vztah (II-38) můžeme interpretovat také takto:

$$\hat{\mathbf{p}}(t+1) = \hat{\mathbf{p}}(t) + (\text{vektor korekčních zesílení}) \times (\text{chyba modelu}),$$

z čehož vyplývá, že pokud se výstup modelu shoduje s naměřenou hodnotou, nebudeme měnit parametry modelu. Tento algoritmus ale stále vyžaduje nalezení inverze matice $\mathbf{M}(t+1)$, což se dá obejít s pomocí věty o inverzi matice:

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1},$$

kde se předpokládá, že u všech matic existuje inverze.

Definujeme:

$$\mathbf{P}(t+1) = \mathbf{M}^{-1}(t+1) = [\mathbf{M}(t) + \mathbf{r}(t)\mathbf{r}^T(t)]^{-1} \quad (\text{II-39})$$

a po dosazení $\mathbf{A} = \mathbf{M}(t) = \mathbf{P}^{-1}(t)$, $\mathbf{B} = \mathbf{r}(t)$, $\mathbf{C} = 1$, $\mathbf{D} = \mathbf{r}^T(t)$ a použití věty o inverzi matice dostaneme:

$$\mathbf{P}(t+1) = \mathbf{P}(t) - \mathbf{P}(t)\mathbf{r}(t)[1 + \mathbf{r}^T(t)\mathbf{P}(t)\mathbf{r}(t)]^{-1}\mathbf{r}^T(t)\mathbf{P}(t). \quad (\text{II-40})$$

V každém kroku tak byla inverze matice nahrazena dělením skalárem $1 + \mathbf{r}^T(t)\mathbf{P}(t)\mathbf{r}(t)$. Vektor korekčních zesílení tak může být psán v následující podobě:

$$\mathbf{k}(t+1) = \mathbf{M}^{-1}(t+1)\mathbf{r}(t) = \mathbf{P}(t+1)\mathbf{r}(t) = \mathbf{P}(t)\mathbf{r}(t) \left[1 - \frac{\mathbf{r}^T(t)\mathbf{P}(t)\mathbf{r}(t)}{1 + \mathbf{r}^T(t)\mathbf{P}(t)\mathbf{r}(t)} \right]$$

a po úpravě:

$$\mathbf{k}(t+1) = \frac{\mathbf{P}(t)\mathbf{r}(t)}{1 + \mathbf{r}^T(t)\mathbf{P}(t)\mathbf{r}(t)}. \quad (\text{II-41})$$

Vektor parametrů $\mathbf{p}$ aktualizujeme dle vztahu (II-42):

$$\hat{\mathbf{p}}(t+1) = \hat{\mathbf{p}}(t) + \mathbf{k}(t+1)[y(t+1) - y_m(t+1)], \quad (\text{II-42})$$

u kterého je výstup modelu popsán v(II-43):

$$y_m(t+1) = \mathbf{r}^T(t)\hat{\mathbf{p}}(t). \quad (\text{II-43})$$

Vztahy (II-41) až (II-43) popisují způsob, jakým se upravuje vektor parametrů $\mathbf{p}$ při iteračním výpočtu. Zbývá tedy ještě odpovědět na otázku, jak provést inicializaci iteračního výpočtu v případě, že nemáme k dispozici žádný odhad těchto parametrů. Principiálně bychom měli čekat tak dlouho, až nashromáždíme dostatek dat a lze provést inverzi matice $\mathbf{M}(t)$. Inicializace algoritmu se pak provede tak, že $\mathbf{P}(t) = \mathbf{M}^{-1}(t)$ a $\hat{\mathbf{p}}(t) = \mathbf{P}(t)\mathbf{v}(t)$. Takto se rekurzivní metoda nejmenších čtverců chová identicky jako její nerekurzivní varianta.

Prakticky lze ale inicializaci provést tak, že vektor parametrů $\hat{\mathbf{p}}(0)$ zvolíme libovolně (např. $\hat{\mathbf{p}}(0) = \mathbf{0}$) a za matici $\mathbf{P}(0)$ dosadíme jednotkovou matici násobenou velikým kladným číslem, čímž říkáme, že inicializačním parametrům $\hat{\mathbf{p}}(0)$ dáváme minimální důvěru (váhu).

II.3.3.2 VEKTOR $\mathbf{P}$ MŮŽE KOLÍSAT

V případě, kdy se vektor parametrů $\mathbf{p}$ pomalu mění, např. protože lineární model byl použit k popisu nelineární soustavy v blízkosti pomalu se posouvajících pracovních bodů, je vhodné použít metodu nejmenších čtverců se zapomínáním. To znamená, že hodnoty naměřené na systému v minulosti budou ovlivňovat aktuální hodnotu parametrů $\mathbf{p}$ s menší váhou než hodnoty aktuální. Pro tento účel byla vyvinuta metoda exponenciálního zapomínání, která ohodnocuje významnost naměřených hodnot exponenciální

váhou, snižující se s uplynutým časem. Aktuální hodnoty jsou ohodnoceny jednotkovou váhou, minulé hodnoty v čase $t+1-n$ jsou váženy koeficientem λ^n . λ označuje koeficient exponenciálního zapomínání a jeho hodnota se volí v intervalu:

$$0 < \lambda \leq 1.$$

$\lambda=1$ znamená, že metoda nejmenších čtverců pracuje bez zapomínání. Čím menší λ , tím rychleji budou zapomínány minulé hodnoty měření.

Pro zavedení zapomínání musíme upravit vztahy (II-34) a (II-35):

$$\mathbf{M}(t+1) = \lambda \mathbf{M}(t) + \mathbf{r}(t) \mathbf{r}^T(t), \quad (\text{II-44})$$

$$\mathbf{v}(t+1) = \lambda \mathbf{v}(t) + \mathbf{r}(t) y(t+1) \quad (\text{II-45})$$

a dále definujeme:

$$\mathbf{k}(t+1) = \frac{\mathbf{P}(t) \mathbf{r}(t)}{\lambda + \mathbf{r}^T(t) \mathbf{P}(t) \mathbf{r}(t)}, \quad (\text{II-46})$$

$$\mathbf{P}(t+1) = \frac{1}{\lambda} [\mathbf{P}(t) - \mathbf{k}(t+1) \mathbf{r}^T(t) \mathbf{P}(t)], \quad (\text{II-47})$$

$$\mathbf{y}_m(t+1) = \mathbf{r}^T(t) \hat{\mathbf{p}}(t), \quad (\text{II-48})$$

$$\hat{\mathbf{p}}(t+1) = \hat{\mathbf{p}}(t) + \mathbf{k}(t+1) [y(t+1) - \mathbf{y}_m(t+1)]. \quad (\text{II-49})$$

Pozn.: čím menší zvolíme λ , tím vyšší bude schopnost algoritmu rychle se přizpůsobovat novým parametrům, ale platíme za to tím, že parametry mohou být více ovlivněny náhodnými poruchami působícími na systém. Proto se při volbě parametru zapomínání λ dostáváme ke kompromisu a typicky se λ volí větší než 0,95.

II.3.4 Obecné metody optimalizace

V předchozích kapitolách (II.3.2 až II.3.3.2) byla prezentována optimalizační metoda nejmenších čtverců ve více podobách. Omezili jsme se však pouze na případy, kdy účelová funkce je kvadrát chyby modelu a samotný model je lineární v parametrech. Toto by ovšem nepostačilo pro sofistikovanější

identifikaci (např. u nelineárních struktur modelů) a proto se nyní budeme zabývat další velkou skupinou numerických metod hledání optima účelové funkce. Tyto metody lze rozdělit do tří hlavních kategorií. Do první kategorie spadají základní optimalizační algoritmy jako jsou gradientní metody a metody jednorozměrného hledání. Tyto algoritmy tvoří jádro důmyslnějších metod patřících do dalších dvou kategorií. Druhou kategorií tvoří optimalizační algoritmy dnes frekventovaně používané a implementované v mnoha softwarových balících. Jmenujme alespoň metodu konjugovaných směrů, kvazinewtonské metody atd. Třetí kategorií tvoří stále vyvíjené algoritmy zabývající se globální optimalizací.

Numerické optimalizační metody obvykle rozloží původní problém na sekvenci problémů elementárních. Např. vícerozměrná optimalizace bývá rozdělena na sekvenci jednorozměrných optimalizací. A to také bývá problémem těchto numerických metod a důvodem, že jedna metoda se může zdát být lepší (resp. rychlejší) na určitém příkladě v porovnání s jinými, ale na odlišné skupině optimalizačních problémů to již může být naopak. V této sekci se budeme zabývat optimalizačními algoritmy s neomezenými parametry. Většina numerických metod je iteračních a počítají optimální parametry v dalším kroku tak, aby se snížila účelová funkce, což otevírá problém inicializace celého řešení. Tohoto problému se také pokusíme dotknout.

Ještě před tím, než začneme jednotlivé metody rozebírat, zmiňme se o zjednodušení, při kterém nám může podstatně pomoci metoda nejmenších čtverců.

II.3.4.1 ČÁSTEČNĚ LINEÁRNÍ MODEL

Mějme model, jehož výstup lineárně závisí na n_l parametrech vektoru $\mathbf{p}_l$ a nelineárně na n_{nl} parametrech vektoru $\mathbf{p}_{nl}$. Výstup modelu můžeme pak psát:

$$\mathbf{y}_m(\mathbf{p}_l, \mathbf{p}_{nl}) = \mathbf{R}(\mathbf{p}_{nl})\mathbf{p}_l \quad (\text{II-50})$$

a předpokládejme kvadratickou účelovou funkci:

$$J(\mathbf{p}_l, \mathbf{p}_{nl}) = [\mathbf{y} - \mathbf{y}_m(\mathbf{p}_l, \mathbf{p}_{nl})]^T \mathbf{Q} [\mathbf{y} - \mathbf{y}_m(\mathbf{p}_l, \mathbf{p}_{nl})]. \quad (\text{II-51})$$

Pro jakoukoliv hodnotu vektoru $\mathbf{p}_{nl}$ je optimální hodnota vektoru $\mathbf{p}_l$, která minimalizuje kvadratickou funkci J , dána dle metody nejmenších čtverců jako funkce vektoru $\mathbf{p}_{nl}$:

$$\hat{\mathbf{p}}_l(\mathbf{p}_{nl}) = [\mathbf{R}^T(\mathbf{p}_{nl})\mathbf{Q}\mathbf{R}(\mathbf{p}_{nl})]^{-1}\mathbf{R}^T(\mathbf{p}_{nl})\mathbf{Q}\mathbf{y} \quad (\text{II-52})$$

a dosazením $\hat{\mathbf{p}}_l$ za $\mathbf{p}_l$ v účelové funkci dostáváme, že tato funkce závisí pouze na nelineárních parametrech:

$$J(\mathbf{p}_{nl}) = [\mathbf{y} - \mathbf{R}(\mathbf{p}_{nl})\hat{\mathbf{p}}_l(\mathbf{p}_{nl})]^T \mathbf{Q} [\mathbf{y} - \mathbf{R}(\mathbf{p}_{nl})\hat{\mathbf{p}}_l(\mathbf{p}_{nl})]. \quad (\text{II-53})$$

Tato úprava nám umožní redukovat rozměr prostoru hledání optima iterační metodou z $n_l + n_{nl}$ na n_{nl} , což podstatně zjednodušuje a zrychluje úkol optimalizace částečně nelineární funkce. Hledání minima v redukovaném prostoru je rychlejší a způsobuje méně numerických problémů. Navíc potřebujeme nyní inicializovat pouze vektor nelineárních parametrů $\mathbf{p}_{nl}$.

II.3.4.2 JEDNODIMENZIONÁLNÍ OPTIMALIZACE

II.3.4.2.1 Volba prohledávaného intervalu

Prvním krokem při jednodimenzionální optimalizaci je definice počátečního intervalu (a,b) , na kterém se bude hledat optimální parametr p . Pro tento účel můžeme např. vyhodnotit nultou a první derivaci účelové funkce v bodě $p_0=0$ (jestliže nelze derivaci spočítat analyticky, pak může být nahrazena konečnou diferencí). Mohou nastat tři případy:

- první derivace je záporná - to znamená, že hodnotu účelové funkce $J(p)$ můžeme snížit zvětšením parametru p na $p_1 = p_0 + \Delta p$, $\Delta p > 0$. Spočteme $J(p_1)$ a jestliže $J(p_1) \geq J(p_0)$, pak $a = p_0$, $b = p_1$ a funkce J má nejméně jedno lokální minimum na tomto intervalu. Jinak se musíme posunout ještě dále od bodu p_0 . To opakujeme až do té doby, kdy se hodnota účelové funkce J začne zvyšovat. Vždy poslední dvě hodnoty parametru p , na kterých byla spočtena hodnota funkce J , mohou být vzaty za hodnoty (a,b) ,
- první derivace je kladná a je tedy nutné postupovat v záporném směru, tzn. $p_1 = p_0 - \Delta p$, $\Delta p > 0$. Jinak postupujeme jako v předchozím bodu,
- první derivace je nulová - to znamená, že můžeme být v minimu. Pak je vhodné vypočítat hodnotu funkce J ve dvou okolních

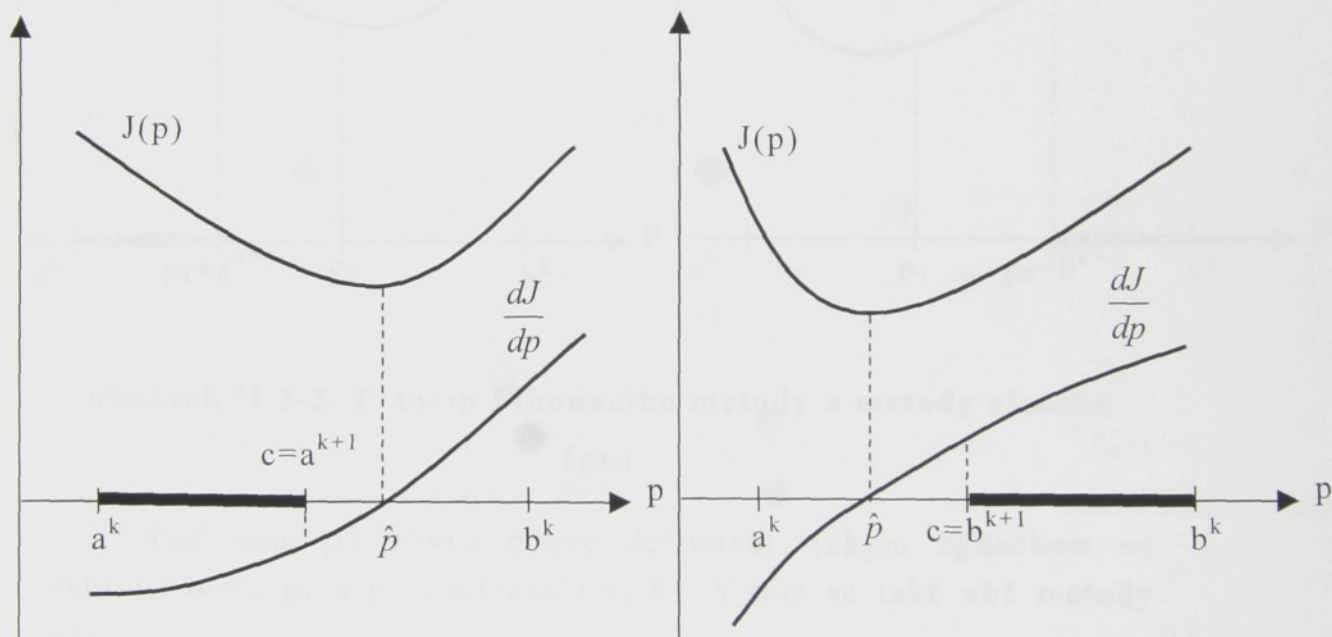
blízkých bodech a tak zjistit více informací o jejím průběhu, jestliže se nejedná o minimum, pak je vhodné se posunout kousek vedle a postupovat dle předchozích bodů.

Jakmile jsme definovali interval prohledávání, můžeme aplikovat různé techniky na jeho zmenšení. Těmi se budeme zabývat v následujících kapitolách (II.3.4.2.2 a II.3.4.2.3).

II.3.4.2.2 Dichotomické vyhledávání

Tato metoda vyhodnocuje hodnotu první derivace účelové funkce J uprostřed intervalu (a,b) . Tuto hodnotu označme $c = \frac{a+b}{2}$.

Jestliže je derivace v tomto bodě kladná, pak položíme $b=c$, jinak $a=c$. V každé iteraci je tudíž interval zmenšen na polovinu.



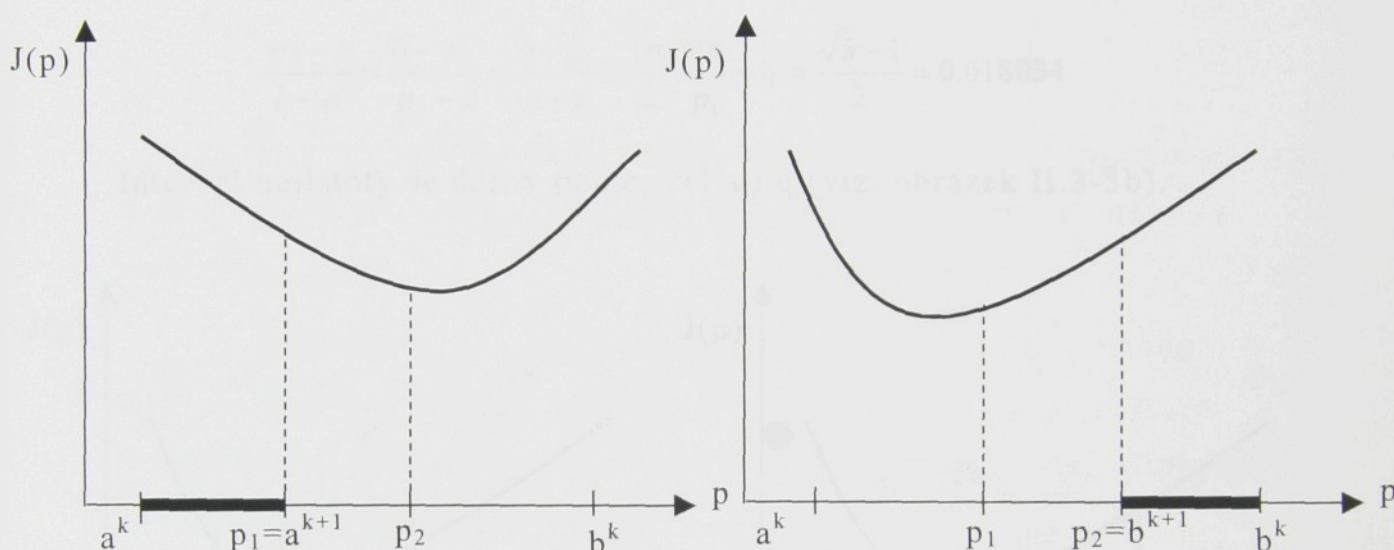
obrázek II.3-1: Dichotomické vyhledávání

I pro velké prohledávané intervaly najde tato metoda minimum s velkou přesností v několika iteracích. Bohužel nám nezaručuje, že vyhledané minimum je globálním a to ani v prohledávaném intervalu.

II.3.4.2.3 Fibonacciho metoda a metoda zlatého řezu

Jestliže neznáme derivaci účelové funkce, popř. by její počítání bylo příliš náročné nebo nepřesné, nemůžeme použít metodu dichotomického vyhledávání. Fibonacciho metoda a metoda zlatého řezu nahrazují výpočet derivace uprostřed prohledávaného intervalu porovnáním dvou hodnot p_1 a p_2 z tohoto intervalu. Interval je tedy rozdělen na tři části a použit následující algoritmus:

jestliže $J(p_2^k) > J(p_1^k)$, pak $a^{k+1} = a^k$ a $b^{k+1} = p_2^k$, jinak $a^{k+1} = p_1^k$ a $b^{k+1} = b^k$ (viz. obrázek II.3-2).



obrázek II.3-2: Princip Fibonacciho metody a metody zlatého řezu

Teď nám již zbývá pouze definovat, jakým způsobem se vybírají body p_1 a p_2 z intervalu (a, b) . V tom se také obě metody liší:

a. Fibonacciho metoda

Fibonacciho posloupnost je definována rekurentním vztahem:

$$F_{i+1} = F_i + F_{i-1} \quad F_0 = F_1 = 1.$$

Výchozí interval rozdělíme v poměru dvou po sobě následujících Fibonacciho čísel F_n a F_{n-1} . Vybereme nový interval tak, aby bod s nejnižší hodnotou účelové funkce ležel uvnitř tohoto intervalu. Výpočtem další jedné hodnoty

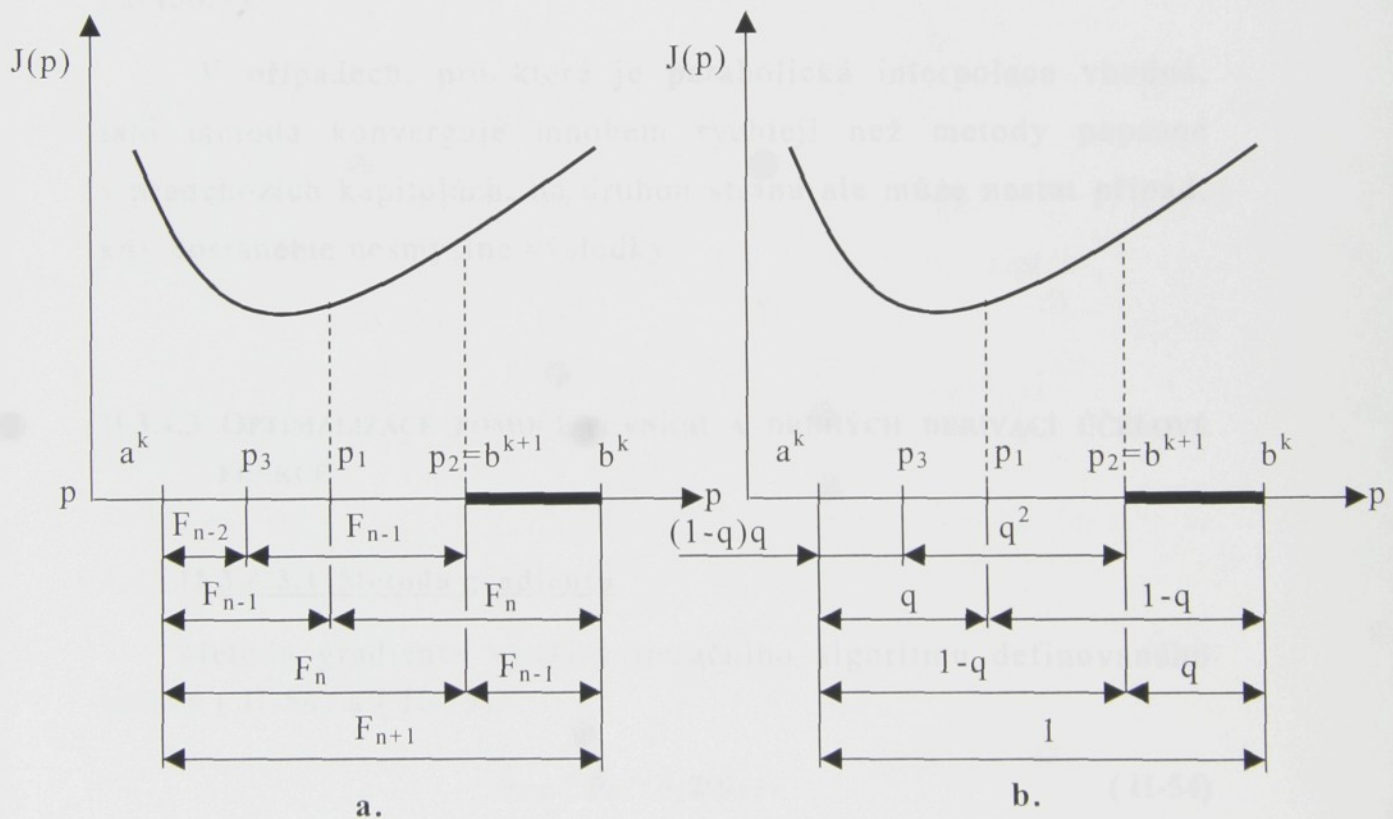
účelové funkce v bodě p_3 rozdělíme tento interval oboustranně v poměru $F_{n-1}:F_{n-2}$ (viz. obrázek II.3-3 a). Postup končí rozdělením posledního intervalu v poměru $F_2:F_1$. Fibonacciho metoda vyžaduje předem daný počet iterací n na rozdíl od metody zlatého řezu. Pro velké n přechází Fibonacciho posloupnost na geometrickou posloupnost s kvocientem zlatého řezu ($q^{-1}=1,618034$).

b. Metoda zlatého řezu

Výchozí interval se rozdělí v poměru zlatého řezu ($q^2 + q - 1 = 0$):

$$\frac{p_2 - a}{b - a} = \frac{b - p_2}{p_2 - a} = \frac{b - p_1}{b - a} = \frac{p_1 - a}{b - p_1} = q = \frac{\sqrt{5} - 1}{2} = 0.618034$$

Interval nejistoty se dělí v poměru $(1-q):q$ (viz. obrázek II.3-3b).



obrázek II.3-3: Ilustrace dělení intervalu, a. Fibonacciho metoda, b. metoda zlatého řezu.

II.3.4.2.4 Parabolická interpolace

V blízkém okolí minima účelové funkce může být tato funkce aproximována parabolou, pro kterou lze analyticky vyjádřit souřadnice vrcholu. Známe-li hodnotu účelové funkce ve třech bodech p_1, p_2, p_3 na prohledávaném intervalu (a,b) , můžeme definovat parabolou, která prokládá tyto tři body takto:

$$P(p) = J(p_1) \frac{(p-p_2)(p-p_3)}{(p_1-p_2)(p_1-p_3)} + J(p_2) \frac{(p-p_1)(p-p_3)}{(p_2-p_1)(p_2-p_3)} + J(p_3) \frac{(p-p_1)(p-p_2)}{(p_3-p_1)(p_3-p_2)},$$

potom hodnota parametru p , která popisuje souřadnici vrcholu paraboly, je dána následujícím vztahem:

$$\hat{p}_p = p_2 - \frac{1}{2} \frac{(p_2-p_1)^2 [J(p_2)-J(p_3)] - (p_2-p_3)^2 [J(p_2)-J(p_1)]}{(p_2-p_1)[J(p_2)-J(p_3)] - (p_2-p_3)[J(p_2)-J(p_1)]}.$$

Tento vrchol paraboly může být použit jako aproximace minima účelové funkce (po prověření, že neodpovídá maximu paraboly).

V případech, pro které je parabolická interpolace vhodná, tato metoda konverguje mnohem rychleji než metody popsané v předchozích kapitolách, na druhou stranu ale může nastat případ, kdy dostaneme nesmyslné výsledky.

II.3.4.3 OPTIMALIZACE POMOCÍ PRVNÍCH A DRUHÝCH DERIVACÍ ÚČELOVÉ FUNKCE

II.3.4.3.1 Metoda gradientu

Metoda gradientu využívá iteračního algoritmu definovaného vztahy (II-54) a (II-55):

$$\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k - \lambda_k \mathbf{g}(\hat{\mathbf{p}}_k), \quad (\text{II-54})$$

$$\text{kde } \mathbf{g}(\hat{\mathbf{p}}_k) = \frac{\partial J}{\partial \mathbf{p}} \Big|_{\mathbf{p}=\hat{\mathbf{p}}_k} \quad (\text{II-55})$$

Algoritmus hledání se ukončí až když je norma gradientu $\|\mathbf{g}(\hat{\mathbf{p}}_{k+1})\|$ dostatečně malá. Zůstává otázka, jak volit délku kroku λ ,

aby počet iterací byl co nejnižší. Existuje více postupů, jak ji měnit, žádný ale není univerzálně nejlepší. Popíšme jeden jednoduchý a efektivní. Při iterování mohou nastat tři případy:

- $J(\hat{p}_{k+1}) < J(\hat{p}_k)$ - to znamená, že vše jde dobře a tak si můžeme dovolit zvětšit krok (např. $\lambda_{k+1} = 1.5\lambda_k$),
- $J(\hat{p}_{k+1}) = J(\hat{p}_k)$ - to většinou znamená, že $\Delta \mathbf{p}$ je příliš malé a tak změna účelové funkce není znatelná v dané přesnosti počítání. Zde také zvětšíme krok λ ,
- a konečně může nastat případ, kdy $J(\hat{p}_{k+1}) > J(\hat{p}_k)$, což znamená, že jsme se dostali příliš daleko a proto krok λ by měl být zmenšen a také je nutno se vrátit o krok zpět tzn. $\hat{p}_{k+1} = \hat{p}_k$.

II.3.4.3.2 Newtonova metoda

Tato metoda je založena na rozvoji účelové funkce v okolí bodu $\hat{p}_k$:

$$J(\hat{\mathbf{p}}_{k+1}) = J(\hat{\mathbf{p}}_k + \Delta \mathbf{p}) = J(\hat{\mathbf{p}}_k) + \mathbf{g}^T(\hat{\mathbf{p}}_k) \Delta \mathbf{p} + \frac{1}{2} \Delta \mathbf{p}^T \mathbf{H}(\hat{\mathbf{p}}_k) \Delta \mathbf{p}, \quad (\text{II-56})$$

kde $\mathbf{g}(\hat{\mathbf{p}}_k)$ je gradient účelové funkce v bodě $\hat{\mathbf{p}}_k$ a $\mathbf{H}(\hat{\mathbf{p}}_k)$ je Hessian v bodě $\hat{\mathbf{p}}_k$ (symetrická $n_p \times n_p$ matice):

$$\mathbf{H}(\hat{\mathbf{p}}_k) = \frac{\partial^2 J}{\partial \mathbf{p} \partial \mathbf{p}^T} \Big|_{\mathbf{p}=\hat{\mathbf{p}}_k}. \quad (\text{II-57})$$

Definujme změnu účelové funkce: $\Delta J = J(\hat{\mathbf{p}}_{k+1}) - J(\hat{\mathbf{p}}_k)$.

Hodnota $\Delta \mathbf{p}$, která vede k největšímu snížení účelové funkce se nalezne vyřešením podmínky optima z (II-58):

$$\frac{\partial \Delta J}{\partial \Delta \mathbf{p}} \Big|_{\Delta \mathbf{p}} = \mathbf{0} \approx \mathbf{H}(\hat{\mathbf{p}}_k) \Delta \mathbf{p} + \mathbf{g}(\hat{\mathbf{p}}_k), \quad (\text{II-58})$$

což znamená krok:

$$\Delta \hat{\mathbf{p}} = -\mathbf{H}^{-1}(\hat{\mathbf{p}}_k) \mathbf{g}(\hat{\mathbf{p}}_k).$$

Za předpokladu, že existuje inverze Hessianu, můžeme Newtonův algoritmus psát následovně:

$$\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k - \mathbf{H}^{-1}(\hat{\mathbf{p}}_k) \mathbf{g}(\hat{\mathbf{p}}_k). \quad (\text{II-59})$$

Prakticky bývá Newtonova metoda většinou implementována ve tvaru s koeficientem λ_k (viz. [HAN1], [MAN], [WAL]):

$$\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k - \lambda_k \mathbf{H}^{-1}(\hat{\mathbf{p}}_k) \mathbf{g}(\hat{\mathbf{p}}_k) \quad (\text{II-60})$$

a inverze Hessovy matice se ekonomičtěji nahrazuje řešením soustavy lineárních rovnic pro $\Delta \mathbf{p}$ ve vztahu (II-61):

$$\mathbf{H}(\hat{\mathbf{p}}_k) \Delta \mathbf{p} = -\lambda_k \mathbf{g}(\hat{\mathbf{p}}_k). \quad (\text{II-61})$$

II.3.4.3.3 Metoda konjugovaných gradientů

Tato metoda předpokládá účelovou funkci jako kvadratickou funkci vektoru parametru $\mathbf{p}$:

$$J(\mathbf{p}) = J(\hat{\mathbf{p}}_k) + \mathbf{g}^T(\hat{\mathbf{p}}_k)(\mathbf{p} - \hat{\mathbf{p}}_k) + \frac{1}{2}(\mathbf{p} - \hat{\mathbf{p}}_k)^T \mathbf{H}(\mathbf{p} - \hat{\mathbf{p}}_k), \quad (\text{II-62})$$

což můžeme psát také ve tvaru:

$$J(\mathbf{p}) = \mathbf{c} + \mathbf{b}^T \mathbf{p} + \frac{1}{2} \mathbf{p}^T \mathbf{H} \mathbf{p} \quad (\text{II-63})$$

se symetrickou Hessovou maticí $\mathbf{H}$. Dále definujeme:

$$\Delta \mathbf{p}_k = \hat{\mathbf{p}}_{k+1} - \hat{\mathbf{p}}_k,$$

$$\mathbf{g}_k = \mathbf{g}(\hat{\mathbf{p}}_k).$$

Jestliže by hodnota $\hat{\mathbf{p}}_{k+1}$ byla vypočítána z $\hat{\mathbf{p}}_k$ jednodimenzionální optimalizací ve směru $\mathbf{d}_k$, pak platí:

$$\Delta \mathbf{p}_k = \lambda \mathbf{d}_k,$$

$$\mathbf{g}_{k+1}^T \mathbf{d}_k = 0.$$

A jestliže by v dalším kroku byla použita Newtonova metoda, pak by bylo splněno:

$$\mathbf{d}_{k+1} = -\mathbf{H}^{-1} \mathbf{g}_{k+1},$$

nebo jinak: $\mathbf{g}_{k+1} = -\mathbf{H} \mathbf{d}_{k+1}.$

Společně tyto dvě metody nabízejí hledat ve směrech, které splňují rovnici (II-64):

$$\mathbf{d}_{k+1}^T \mathbf{H} \mathbf{d}_k = 0. \quad (\text{II-64})$$

Tyto směry jsou označovány jako tzv. konjugované (viz. obrázek II.3-4).

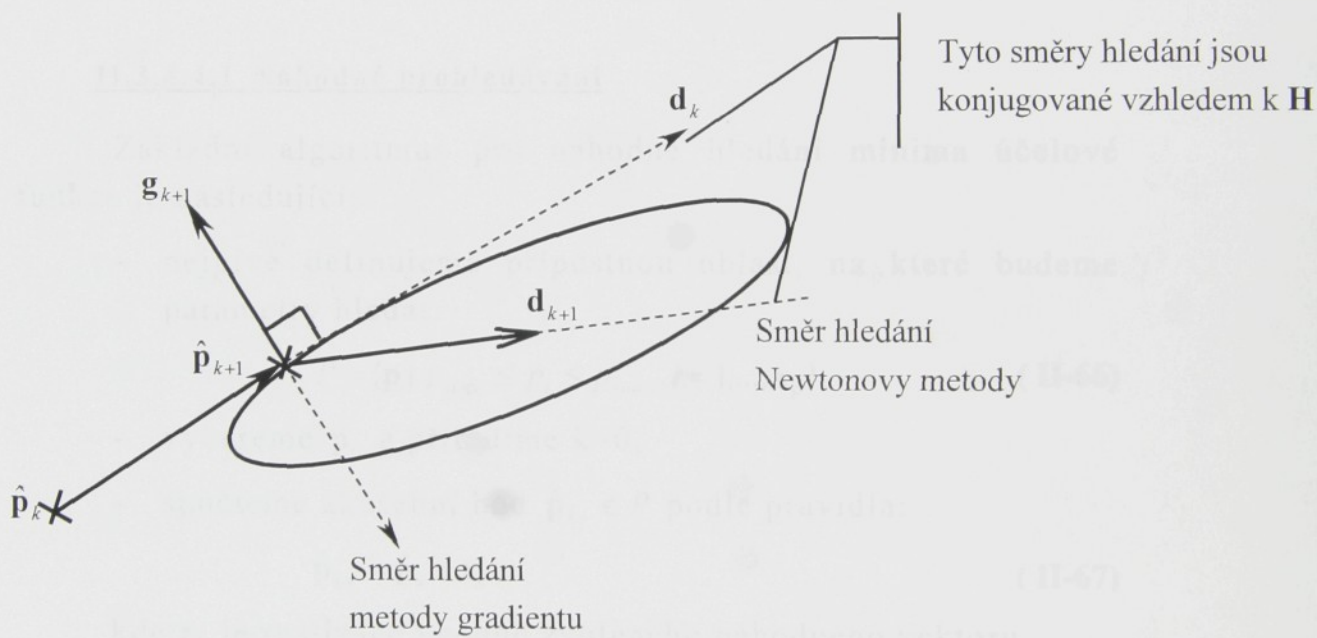
$$\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k + \lambda^* \mathbf{d}_k,$$

$$\lambda^* = \frac{\mathbf{d}_k^T \mathbf{g}(\hat{\mathbf{p}}_k)}{\mathbf{d}_k^T \mathbf{H} \mathbf{d}_k}, \text{ jestliže známe Hessovu matici, jinak:}$$

$$\lambda^* = \arg \min_{\lambda} J(\hat{\mathbf{p}}_k + \lambda \mathbf{d}_k),$$

$$\mathbf{g}_{k+1} = \mathbf{g}(\hat{\mathbf{p}}_{k+1}),$$

$$\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \frac{(\mathbf{g}_{k+1} - \mathbf{g}_k)^T \mathbf{g}_{k+1}}{\mathbf{g}_k^T \mathbf{g}_k} \mathbf{d}_k.$$



obrázek II.3-4: Konjugované směry hledání

II.3.4.4 GLOBÁLNÍ OPTIMALIZACE

Cílem globální optimalizace je nalezení nejlepší možné hodnoty účelové funkce J v odpovídajícím bodě $\tilde{\mathbf{p}}$ tak, že pro všechny možné hodnoty parametrů $\mathbf{p}$ platí:

$$J(\mathbf{p}) \geq J(\tilde{\mathbf{p}}). \quad (\text{II-65})$$

Dnes již existuje mnoho globálních optimalizačních metod. Některé z nich jsou velmi komplexní pro dvojdimenzionální typy úloh, ale zdají se být těžce zobecnitelné na vícedimenzionální problémy. V této práci se podíváme pouze na jednu metodu. Tato metoda, založená na náhodném prohledávání, je jednoduchá, ale nezaručuje nalezení globálního minima. Dnes již sice existují i metody zaručující výsledky (např. deterministická optimalizace [WAL]), ale za cenu větší složitosti a ne pro obecnou účelovou funkci.

II.3.4.4.1 Náhodné prohledávání

Základní algoritmus pro náhodné hledání minima účelové funkce je následující:

- nejprve definujeme přípustnou oblast, na které budeme parametry hledat:

$$P = \{\mathbf{p} \mid p_{i\min} \leq p_i \leq p_{i\max}, i = 1, \dots, n_p\} \quad (\text{II-66})$$

- vybereme $\hat{\mathbf{p}}_0$ a přiřadíme $k=0$,
- spočteme zkušební bod $\hat{\mathbf{p}}_{k+} \in P$ podle pravidla:

$$\hat{\mathbf{p}}_{k+} = \hat{\mathbf{p}}_k + \mathbf{r}_k, \quad (\text{II-67})$$

kde $\mathbf{r}_k$ je realizace vhodně zvoleného náhodného vektoru

- jestliže $J(\hat{\mathbf{p}}_{k+}) < J(\hat{\mathbf{p}}_k)$ pak $\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_{k+}$, jinak $\hat{\mathbf{p}}_{k+1} = \hat{\mathbf{p}}_k$,
- nakonec zvětšíme o jedničku k a algoritmus opakujeme od výběru nového zkušebního bodu.

Jestliže nemáme žádné preference jak vybrat startovací bod $\hat{\mathbf{p}}_0$, pak se vybírá ve středu přípustné oblasti P .

Posunutí do dalšího testovaného bodu $\mathbf{r}_k$ je možné volit jako náhodně generovaný vektor dle Gaussova rozložení $N(\mathbf{0}, \Sigma(\sigma))$ s nulovou střední hodnotou a rozptylem $\Sigma(\sigma)$ ořezaný tak, aby vektor parametrů patřil do přípustné oblasti P (II-66).

Jednoduchá a účinná modifikace algoritmu je měnit rozptyl náhodného rozložení v průběhu prohledávání. Pro začátek můžeme např. zvolit:

$$^1\sigma = p_{\max} - p_{\min}, \quad (\text{II-68})$$

což znamená velké posuny vektoru p . S přibývajícemi iteracemi se rozptyl zjemňuje dle vztahu (II-69):

$$^i\sigma = \frac{^{i-1}\sigma}{10}, i = 2, 3, \dots, 5, \quad (\text{II-69})$$

pro jemnější a jemnější prohledávání přípustné oblasti. Každému $^i\sigma$ rozptylu může např. být přiřazeno $100/i$ iterací, čímž přiřadíme více pokusů větším skokům parametrů.

Algoritmus většinou končí vyčerpáním maximálního množství iterací nebo dosažením minimálního požadovaného rozptylu.

II.4 VYHODNOCENÍ KVALITY MODELU

Po nalezení modelu vyvstává otázka o tom, jestli je tento model dostatečně dobrý a použitelný pro zamýšlený účel či ne. Zodpovězení této otázky je vždy z velké míry subjektivní a také velmi obtížné. Model je vhodné konfrontovat se všemi dostupnými informacemi, předpoklady o systému počínaje, změřenými daty a zkušenostmi s používáním modelu a systému konče ([WAL]). Všeobecnou radou pro zvýšení důvěry k modelu je vyzkoušet tak mnoho hodnotících kritérií kvality jak jen je to možné.

Prvním představitelným prostředkem testování kvality je použití zdravého rozumu; např. jestliže estimovaný parametr reprezentuje délku tyče, musí být přinejmenším kladný. Takové testy připadají do kategorie předpokladů o systému a jsou důležité zejména u modelů, u nichž mají parametry fyzikální význam.

Naprostá většina metod vyhodnocování kvality modelu je založena na změřených datech. Základním testem je vyšetření rozptylu parametrů. Veliký rozptyl v porovnání s nalezenými parametry znamená, že něco pravděpodobně není v pořádku. Další docela užitečný test je nalezení několika modelů paralelně. Např. můžeme porovnat frekvenční charakteristiku parametrického a neparametrického modelu.

Nejvšestrannějším a nejvyužívanějším nástrojem vyhodnocování kvality modelu je simulace. Studovaný systém a získaný model jsou buzeny stejným vstupním signálem a následovně jsou porovnávány jejich výstupy. Pro nestranné porovnání je žádoucí, aby vyhodnocení kvality modelu bylo založeno na čerstvých datech, tedy ne na datech, která byla použita pro nalezení modelu.

Často také stojí za to zkoumat sekvenci residuí, zvláště pak na čerstvých datech. Velmi jednoduchým testem je zobrazení sekvence residuí a intuitivní zkoumání, jestli mají předpokládané vlastnosti. Jako doplněk k tomuto testu se často zkoumá předpokládaná "whiteness" (jak moc se podobají bílému šumu) sekvence residuí zobrazením autokovarianční funkce (II-70):

$$R(\tau T) = \frac{1}{N} \sum_{k=1}^{N-\tau} e(kT, \mathbf{p}) e(kT + \tau T, \mathbf{p}), \quad (\text{II-70})$$

která má být pro $\tau \neq 0$ malá. Navíc předpokládána nezávislost mezi vstupním signálem a sekvencí residuí se testuje zobrazením kroskovarianční funkce mezi residuí a budícím signálem:

$$R(\tau T) = \frac{1}{N} \sum_{k=\max(1,1-\tau)}^{\min(N,N-\tau)} e(kT, \mathbf{p}) u(kT + \tau T), \quad (\text{II-71})$$

pro různá τ . Tato funkce by měla být také malá, jinak to znamená, že jsme v modelu nepostihli všechny vlivy vstupu systému na výstup. Jinak řečeno, v systému se stále skrývá nenamodelovaná dynamika.

Nakonec bychom měli zdůraznit, že čím více testům podrobíme model, tím je větší šance vymezení hranic použitelnosti a "pravdivosti" získaného modelu.

V této kapitole byly zmíněny pouze nejpoužívanější metody ověření platnosti modelu. V případě dalšího zájmu doporučujeme literaturu zabývající se tímto problémem (např.: [LJU2], [LJU5], [LJU6], [SOD]).

III Identifikace ve spojité časové oblasti

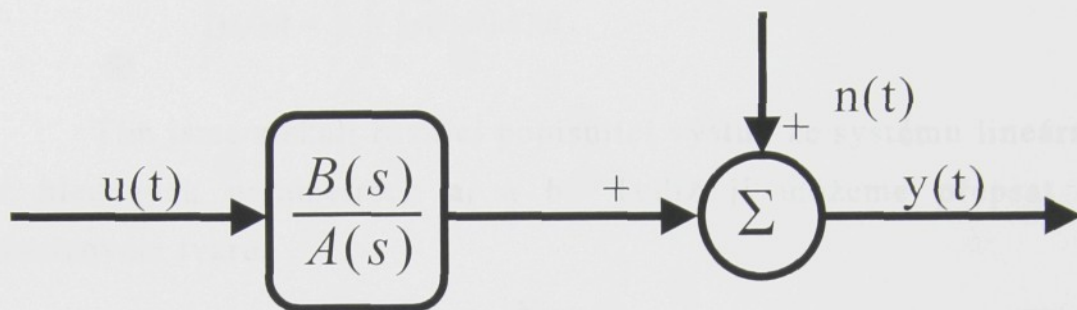
Tato kapitola, která je vlastním těžištěm předkládané práce, se zabývá metodami hledání modelů dynamických systémů z experimentální dat ve spojité časové oblasti. Po stručném úvodu (kapitola III.1) se nejdříve věnujeme metodám identifikace struktur modelů lineárních v parametrech. Pro tuto skupinu jsou zde prezentovány dvě metody: integrační a derivační (kapitoly III.2 a III.3). Rozšíření těchto metod na soustavy s nenulovými počátečními podmínkami je charakterizováno v kapitole III.4. Část III.5 je věnována způsobu použití těchto metod na některé nelineární struktury modelů. Optimální parametry nelineárních diferencovatelných struktur modelů dokáže nalézt metoda kvazilinearizace (kapitola III.6). Způsobem hledání optimálních parametrů nejsložitějšího případu, nediferencovatelnými nelineárními strukturami modelů, se zabývá kapitola III.8. A konečně v závěru (kapitola III.9) je rozebrán vliv poruch na odhad parametrů derivační a integrační metodou.

III.1 ÚVOD

V historii vývoje identifikačních metod se spojité časové oblasti v posledních desetiletích věnovalo velmi málo úsilí a v literatuře jsou jen spoře popsány některé metody hledání spojitých modelů. Ačkoliv počátky vývoje teorie identifikace jsou spojeny se spojitými modely (připomeňme metodu prof. Strejce), tyto byly velmi brzy nahrazeny modely diskrétními, které pracují s daty tak jak jsou dnes obvykle snímána tzn. v diskrétní formě. Tyto metody zaznamenaly v poslední době ohromný rozmach a zcela zastínily algoritmy hledající modely ve spojité časové oblasti. Naproti tomu, s rostoucí rychlostí zpracování dat současnými počítači jsme schopni snímat data s tak malou vzorkovací periodou, že takto můžeme nahradit spojitý signál pro většinu mechatronických technologií (za předpokladu, že vzorkovací perioda je několikanásobně menší než nejnižší časová konstanta systému). A i když zpracování takového velkého množství dat klade nároky na výkon počítače (popř. na čas výpočtu) můžeme zároveň těžit z výhod, které nám poskytne model systému ve spojité časové oblasti. Zmiňme se alespoň o některých v porovnání s modely v diskrétní oblasti. Tyto výhody vycházejí z jednoduchého principu: studovaný systém se obvykle nachází ve spojité časové oblasti a proto je výhodné ho i tak popisovat. Naproti tomu výstupy diskrétních modelů jsou počítány pouze v určitých časových okamžicích odpovídajících vzorkování a tím nás trochu svazují. Z toho plyne, že výstup diskrétního modelu v sobě může skrývat oscilace odpovídajícího spojitého modelu. A konečně parametry diskrétních modelů nemají všeobecně jasný fyzikální význam (parametry získané diskretizací závisí na volbě vzorkovací periody) a ani vlastnosti diskrétních a spojitých modelů nejsou vždy analogické. Z toho také vyplývá, že sestavování spojitých modelů z diskrétních může být složité a nepřesné. Proto se v této kapitole pokusíme představit metody identifikace modelů ve spojité oblasti.

III.2 INTEGRAČNÍ METODA

Pro jednoduchost zpočátku uvažujme jednorozměrný systém, jehož výstup lze popsat lineární diferenciální rovnicí v závislosti na vstupech a vnitřních proměnných. Dále připouštíme, že na výstup ze systému může působit porucha, kterou zase pro jednoduchost nejprve uvažujme v podobě bílého šumu (obrázek III.2-1):



obrázek III.2-1: Lineární jednorozměrný model

Známo: $u(t)$, $y(t)$.

Neznámo: $B(s)$, $A(s)$, $n(t)$.

Zjišťujeme: $B(s)$, $A(s)$,

kde s ...je Laplaceův operátor,

na ...je stupeň polynomu $A(s)$ viz. (III-1),

nb ...je stupeň polynomu $B(s)$ viz. (III-2).

$$A(s) = a_0 + a_1s + a_2s^2 + \dots + a_{na-1}s^{na-1} + s^{na}, \quad (\text{III-1})$$

$$B(s) = b_0 + b_1s + b_2s^2 + \dots + a_{nb}s^{nb}. \quad (\text{III-2})$$

Diferenciální rovnici popisující chování takového systému lze psát ve tvaru:

$$y^{(na)}(t) + \sum_{i=0}^{na-1} a_i y^{(i)}(t) = \sum_{i=0}^{nb} b_i u^{(i)}(t) \quad (\text{III-3})$$

a tu pak jednoduše upravíme na (III-4):

$$y(t)^{(na)} = - \sum_{i=0}^{na-1} a_i y^{(i)}(t) + \sum_{i=0}^{nb} b_i u^{(i)}(t) \quad (\text{III-4})$$

a po na integracích (III-4) dostaneme (III-5):

$$y(t) = - \sum_{i=0}^{na-1} a_i \int_{(na-i)} y(t) dt + \sum_{i=0}^{nb} b_i \int_{(na-i)} u(t) dt, \quad (\text{III-5})$$

kde $\int_{(i)}$ je označení i -tých integrálů, takže např.

$$\int_3 y(t) dt = \int_0^{t_f} \left(\int_0^{t_f} \left(\int_0^{t_f} y(t) dt \right) dt \right) dt.$$

Tím jsme získali rovnici popisující výstup ze systému lineární v hledaných parametrech a_i a b_i . Tudíž ji můžeme přepsat v maticovém tvaru:

$$\mathbf{y}_m = \mathbf{X} \mathbf{p}, \quad (\text{III-6})$$

kde

$$\mathbf{X} = \left[\int_{na} u dt, \int_{na-1} u dt, \dots, \int_{na} y dt, - \int_{na-1} y dt, \dots, - \int y dt \right],$$

$$\mathbf{p} = [b_0, b_1, \dots, b_{nb} \mid a_0, a_1, \dots, a_{na-1}]^T, \quad \int y dt = \begin{bmatrix} \int_0^T y(t) dt & \int_0^{2T} y(t) dt & \dots & \int_0^{NT} y(t) dt \end{bmatrix}^T.$$

Jestliže pro nalezení optimálních parametrů ze změřených signálů $u(t)$ a $y(t)$ použijeme metodu nejmenších čtverců, kde minimalizovaná účelová funkce je definována jako kvadrát rozdílu odchylek měření a modelu viz. (III-7):

$$J = \mathbf{e}^T \mathbf{e} \rightarrow \min, \quad (\text{III-7})$$

$$\text{kde } \mathbf{e} = \mathbf{y} - \mathbf{y}_m.$$

Minimalizací účelové funkce (III-7) získáme vztah (III-8) pro optimální parametry:

$$\left. \frac{\partial J}{\partial \mathbf{p}} \right|_{\mathbf{p}=\hat{\mathbf{p}}} = -\frac{1}{2} \mathbf{X}^T (\mathbf{y} - \mathbf{X} \hat{\mathbf{p}}) = 0 \quad (\text{III-8})$$

a za předpokladu, že existuje inverze matice $\mathbf{X}^T \mathbf{X}$ můžeme psát vektor parametrů následovně (III-9):

$$\hat{\mathbf{p}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (\text{III-9})$$

Jak již bylo naznačeno v předchozích kapitolách, v praxi se model většinou nehledá pomocí inverze matice $\mathbf{X}^T \mathbf{X}$ kvůli špatným numerickým vlastnostem této metody. Proto se velmi často používají metody dekompozice matice $\mathbf{X}$ (jmenujme alespoň Choleského dekompozici, QR dekompozici...[PRE], [NIU]).

Pro tento případ můžeme QR dekompozici popsat následovně:

- matici $\mathbf{X}$ rozložíme tak, že:

$$\mathbf{X} = \mathbf{Q} * [\mathbf{R} \quad \mathbf{0}]^T \text{ a } \mathbf{Q}^T \mathbf{y} = [\mathbf{z}_1 \quad \mathbf{z}_2]^T,$$

kde $\mathbf{R}$ je horní trojúhelníková matice a $\mathbf{Q}$ je taková čtvercová matice, že platí:

$$\mathbf{Q}^T \cdot \mathbf{Q} = \mathbf{E}$$

a poté vyřešením soustavy rovnic $\mathbf{R}\hat{\mathbf{p}} = \mathbf{z}_1$ prostým dosazením od posledního řádku nahoru získáme optimální parametry identifikovaného modelu.

Pozn. Nutno podotknout, že takto nalezené parametry nemusí být optimální z hlediska globálního minima účelové funkce, protože ve výpočtu nahrazujeme vnitřní proměnné modelu změřenými. To se nám negativně projeví hlavně pokud se tyto od sebe podstatně liší (např. při špatné volbě struktury modelu). I přes to ale lze tuto metodu doporučit pro její numerickou nenáročnost a dosažené výsledky je možné např. použít jako inicializační odhad následujících iteračních optimalizačních algoritmů.

Ilustrační Příklad:

Uvažujme následující příklad: mějme naměřenou (získanou simulací) přechodovou charakteristiku (známo: $\mathbf{u}(t_i)$, $\mathbf{y}(t_i)$) soustavy druhého řádu popsané diferenciální rovnicí ve tvaru:

$$5u = \ddot{y} + 1\dot{y} + 1y,$$

což přepíšeme na dif. rovnici s hledanými parametry:

$$\ddot{y} = b_0 u + a_1 \dot{y} + a_0 y.$$

Model uvažujme ve tvaru:

$$\mathbf{y}_m = \mathbf{X}\mathbf{p},$$

kde:

$$\mathbf{X} = \left[\int_2 \mathbf{u} dt \mid - \int_2 \mathbf{y} dt, - \int_1 \mathbf{y} dt \right],$$

$$\mathbf{p} = [b_0, a_0, a_1]^T.$$

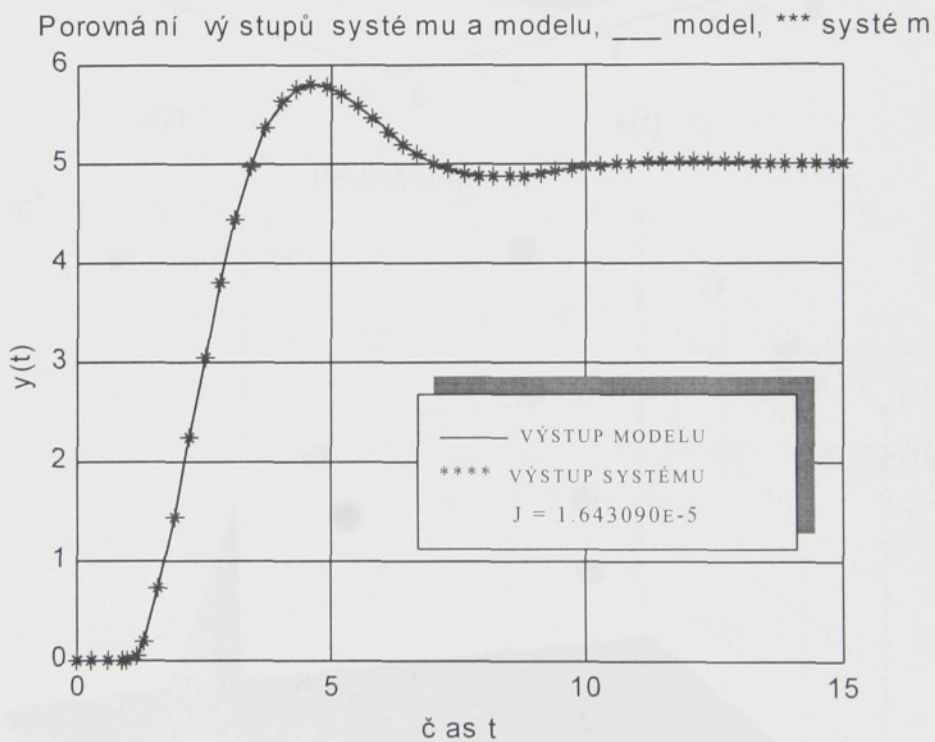
Po provedené optimalizaci metodou nejmenších čtverců jsme našli optimální parametry takto:

$$p(1) = 5.0468, p(2) = 1.0035, p(3) = 1.0093.$$

Odchyłky těchto parametrů od skutečných jsou nejvíce způsobeny přesností numerické integrace, kterou lze zpřesnit zmenšením vzorkování. Po zmenšení periody vzorkování na polovinu jsme obdrželi tyto výsledky:

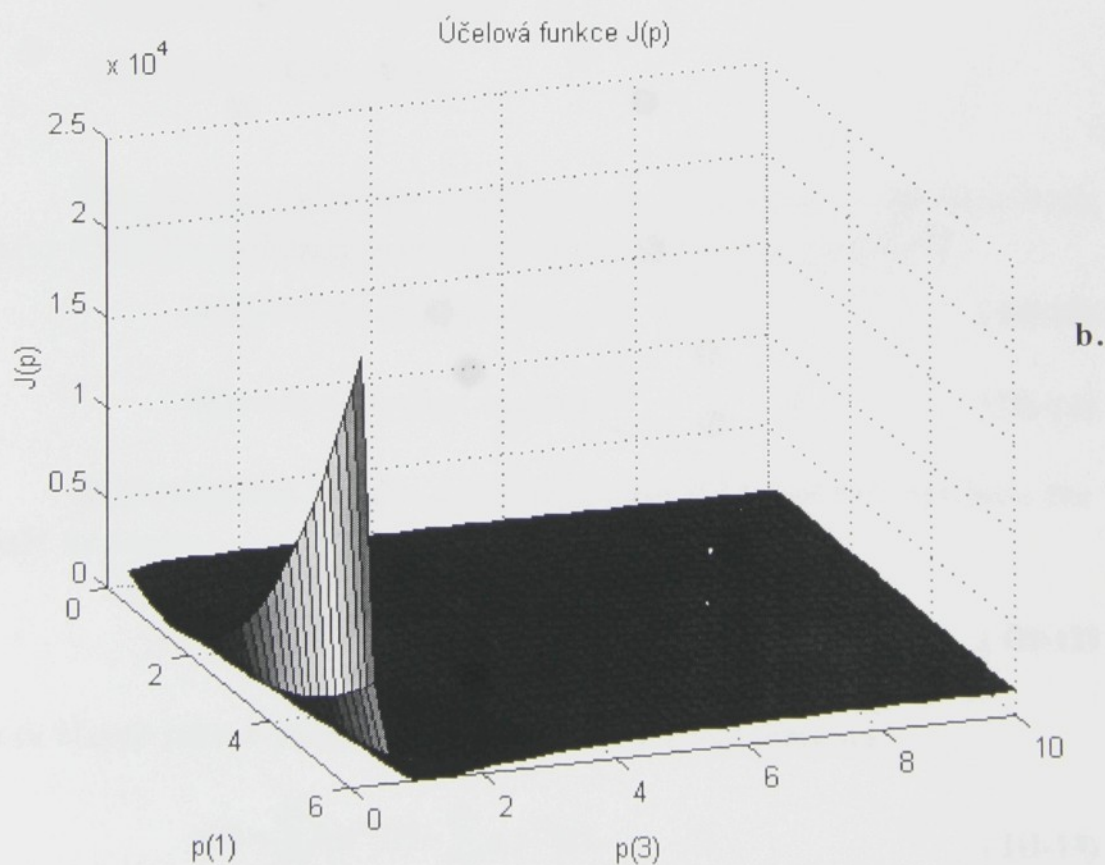
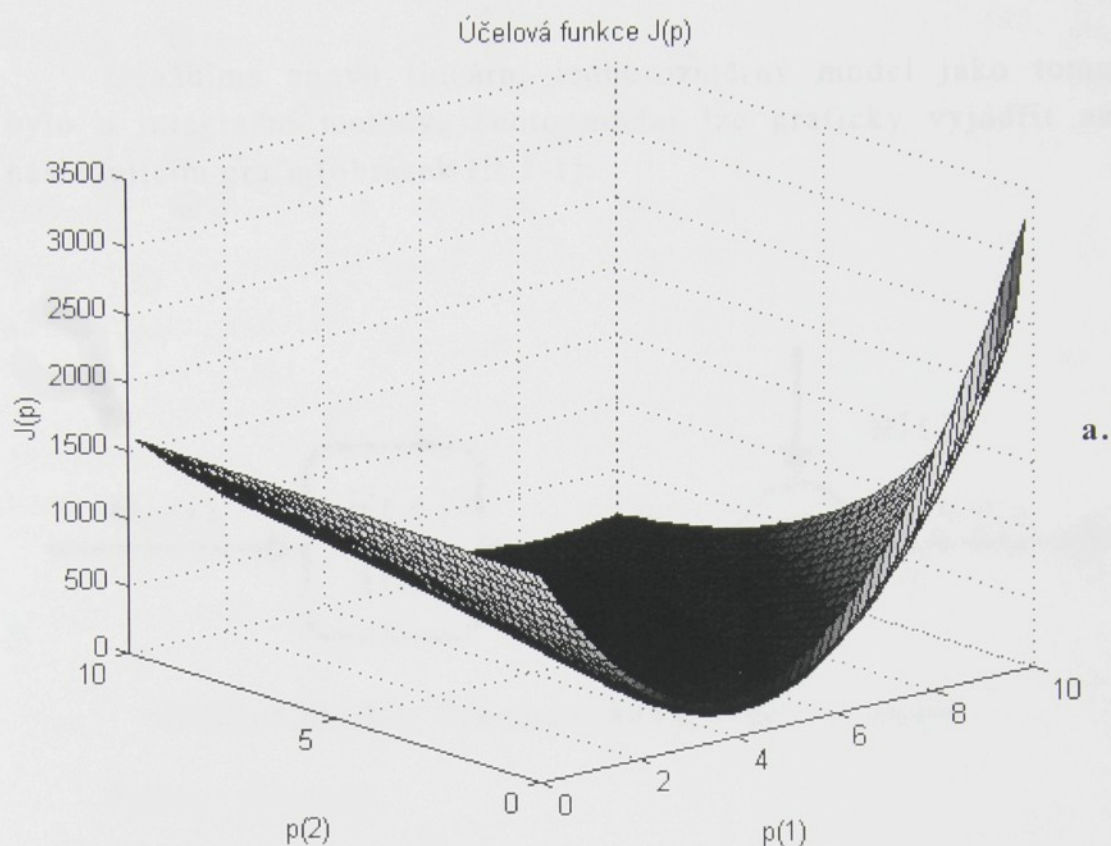
$$p(1) = 5.0022, p(2) = 1.0002, p(3) = 1.0004.$$

Na následujícím obrázku je pak znázorněno porovnání výstupů modelu a systému vybuzeých skokovým signálem.



obrázek III.2-2: Porovnání výstupů modelu a systému

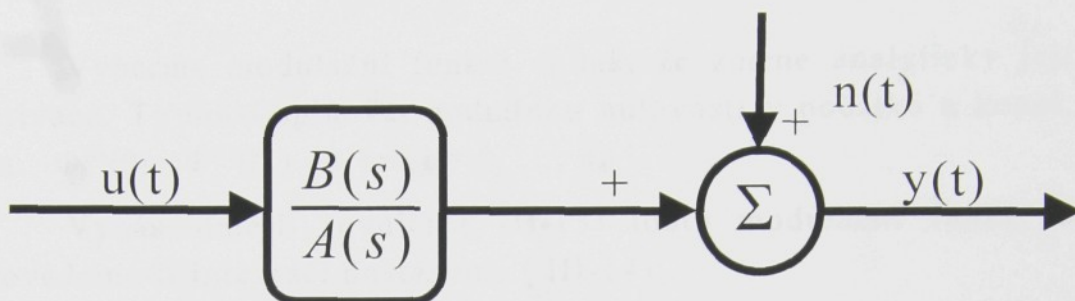
Na následujícím grafu (obrázek III.2-3) je znázorněn průběh účelové kvadratické funkce $J(\mathbf{p})$. Všimněme si, že účelová funkce je hladká a jednotlivé parametry mají na její průběh velice nerovnoměrný vliv.



obrázek III.2-3: Účelová funkce, a. $p(3)=1$, $p(1), p(2)$ se mění, b. $p(2)=1$, $p(1), p(3)$ se mění

III.3 DERIVAČNÍ METODA

Uvažujme znovu lineární jednorozměrný model jako tomu bylo u integrační metody. Tento model lze graficky vyjádřit na následujícím grafu (obrázek III.3-1):



obrázek III.3-1: Lineární jednorozměrný model

Známo: $u(t)$, $y(t)$.

Neznámo: $B(s)$, $A(s)$, $n(t)$.

Zjišťujeme: $B(s)$, $A(s)$.

Pro změnu, bez újmy na obecnosti, budeme uvažovat polynom $A(s)$ v trošku odlišném tvaru od (III-1), $B(s)$ zůstává stejný:

$$A(s) = 1 + a_1 s + a_2 s^2 + \dots + a_{na-1} s^{na-1} + a_{na} s^{na}, \quad (\text{III-10})$$

$$B(s) = b_0 + b_1 s + b_2 s^2 + \dots + a_{nb} s^{nb}. \quad (\text{III-11})$$

Diferenciální rovnici popisující chování takového systému lze psát ve tvaru:

$$y(t) + \sum_{i=1}^{na} a_i y^{(i)}(t) = \sum_{i=0}^{nb} b_i u^{(i)}(t) \quad (\text{III-12})$$

a tu stejně jako v předchozí kapitole III.2 prepíšeme na:

$$y(t) = \sum_{i=0}^{nb} b_i u^{(i)}(t) - \sum_{i=1}^{na} a_i y^{(i)}(t), \quad (\text{III-13})$$

čímž jsme dostali rovnici popisující výstup ze systému v závislosti na derivacích vstupu a výstupu. Samozřejmě by bylo možné tyto

derivace numericky vypočítat a přímo použít v modelu. To je ale ve většině případů krajně nevhodné, protože vysokofrekvenční šum, který nám ve většině případů zhoršuje měření, by se po několikanásobné derivaci drasticky zvětšil a tím nám zastínil užitečné informace o systému. Proto použijeme přístup tzv. *modulačních funkcí*, jehož princip se pokusíme objasnit na následujících řádcích.

Vyberme modulační funkci Φ tak, že známe analyticky její derivace. Ty musí splňovat podmínku nulovosti v počátku a konci, tzn. $\Phi^{(i)}(0) = \Phi^{(i)}(t_f) = 0$ pro $i = 0, \dots, n_a - 1$.

Vynásobíme-li rovnici (III-13) touto modulační funkcí a provedeme-li integraci dostaneme (III-14):

$$\int_0^{t_f} y(t)\Phi(t)dt = \sum_{i=0}^{nb} b_i \int_0^{t_f} u^{(i)}(t)\Phi(t)dt - \sum_{i=1}^{na} a_i \int_0^{t_f} y^{(i)}(t)\Phi(t)dt. \quad (\text{III-14})$$

Rovnici (III-14) lze poté za předpokladů o modulační funkci přepsat na rovnici (III-15):

$$\int_0^{t_f} y(t)\Phi(t)dt = \sum_{i=0}^{nb} b_i (-1)^i \int_0^{t_f} u(t)\Phi^{(i)}(t)dt - \sum_{i=1}^{na} a_i (-1)^i \int_0^{t_f} y(t)\Phi^{(i)}(t)dt. \quad (\text{III-15})$$

A zde vidíme cíl celého snažení: v rovnici (III-15) jsme se zbavili derivací měřeného signálu. Ty jsou nahrazeny analyticky známými derivacemi modulační funkce.

Vysvětlení:

a. Řád derivace 1:

$$(y\Phi)' = y'\Phi + y\Phi'$$

$$\int_0^{t_f} y'\Phi = [y\Phi]_0^{t_f} - \int_0^{t_f} y\Phi' \quad \text{a protože } \Phi(0) = \Phi(t_f) = 0 \quad \Rightarrow \quad \int_0^{t_f} y'\Phi = - \int_0^{t_f} y\Phi'.$$

b. Řád derivace 2:

$$\text{analogicky můžeme psát: } \int_0^{t_f} y''\Phi = [y'\Phi]_0^{t_f} - \int_0^{t_f} y'\Phi',$$

$$\text{kde } [y'\Phi]_0^{t_f} = 0 \text{ a protože také } \Phi'(0) = \Phi'(t_f) = 0 \Rightarrow \int_0^{t_f} y'\Phi' = -\int_0^{t_f} y\Phi''$$

$$\Rightarrow \int_0^{t_f} y''\Phi = \int_0^{t_f} y\Phi''$$

atd. můžeme postupovat do vyšších řádů.

$$*\text{Pozn. } y' = \frac{dy(t)}{dt} \quad y'' = \frac{d^2 y(t)}{dt^2}.$$

Konec vysvětlení.

Protože data jsou ve skutečnosti snímána diskrétně, lze rovnici (10) přepsat ve vektorovém tvaru:

$$\langle y | \Phi \rangle = \sum_{i=0}^{nb} b_i (-1)^i \langle u | \Phi^{(i)} \rangle - \sum_{i=1}^{na} a_i (-1)^i \langle y | \Phi^{(i)} \rangle, \quad (\text{III-16})$$

kde $\langle y | \Phi \rangle$ značí skalární součin.

Nevýhodou této metody je, že volba modulační funkce není triviální a pro zjištění parametrů modelu potřebujeme nejméně tolik nezávislých testovacích funkcí, kolik máme neznámých parametrů modelu. Jeden z možných postupů sestavení modulační funkce si předvedeme v následujícím odstavci.

Pro potřeby minimalizace metodou nejmenších čtverců sestavíme matici R takto:

$$\mathbf{X}(k,:) = [\langle u | \Phi_k \rangle, -\langle u | \Phi_k' \rangle, \langle u | \Phi_k'' \rangle, \dots, (-1)^{nb} \langle u | \Phi_k^{(nb)} \rangle, \langle y | \Phi_k \rangle, -\langle y | \Phi_k' \rangle, \langle y | \Phi_k'' \rangle, \dots, (-1)^{na+1} \langle y | \Phi_k^{(na)} \rangle], \quad (\text{III-17})$$

kde: $\mathbf{X}(k,:)$ značí k -tý řádek matice $\mathbf{X}$,

Φ_k ... k -tá lineárně nezávislá modulační funkce.

Vektor hledaných parametrů $\mathbf{p}$ sestavíme dle (III-18):

$$\mathbf{p} = [b_0, b_1, \dots, b_{nb} | a_1, a_2, \dots, a_{na}]^T. \quad (\text{III-18})$$

A minimalizuje metodou nejmenších čtverců rovnici (III-19):

$$\mathbf{Y} = \mathbf{Xp}, \quad (\text{III-19})$$

ve které jsou řádky vektoru $\mathbf{Y}$ dány následovně:

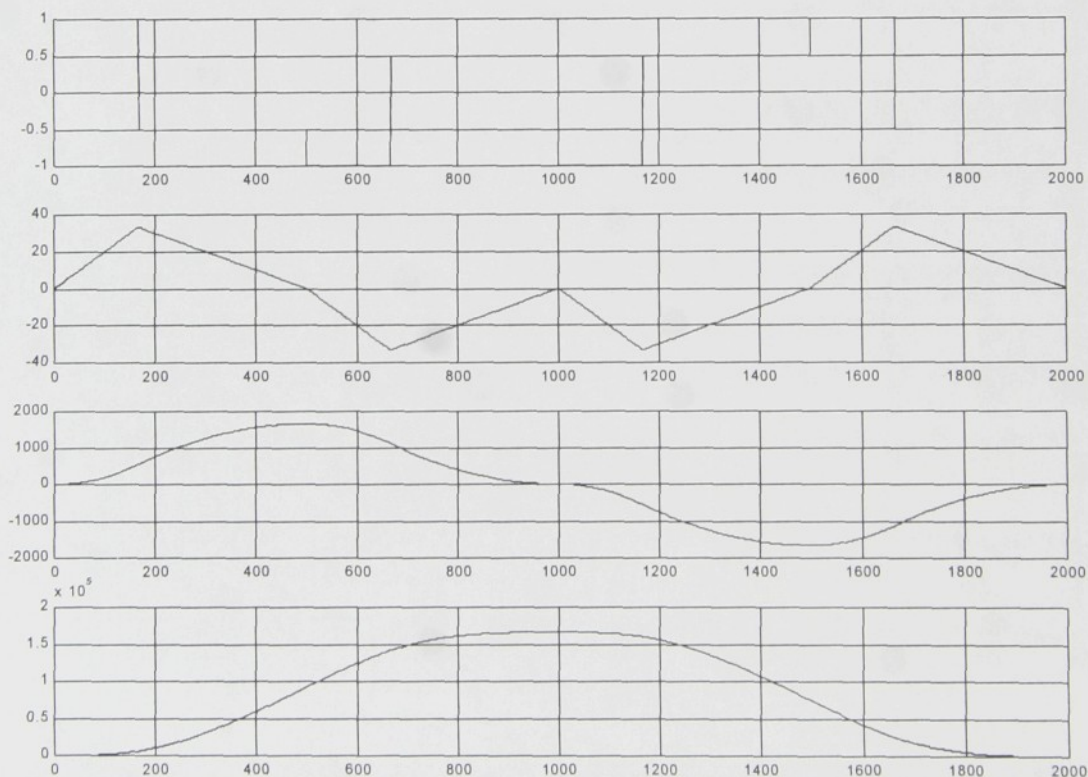
$$y(k) = [\langle y | \Phi_k \rangle], \text{ kde } y(k) \dots \text{ je } k\text{-tý prvek vektoru } \mathbf{Y}.$$

Dále již pak při minimalizaci postupujeme standardními kroky metody nejmenších čtverců stejně jako u integrační metody a nalezneme vektor optimálních parametrů lineárního modelu.

Generování lineárně nezávislých modulačních funkcí:

Zmiňme se o jednom možném způsobu vytváření modulačních funkcí. Při generování modulačních funkcí postupujeme v opačném sledu, tzn. že nejdříve vytvoříme nejvyšší derivaci této funkce:

-ta je po částech konstantní a skládá se z 2^n úseků. Tyto úseky jsou generovány jako řada vznikající zrcadlovým otočením předcházející posloupnosti. Základem řady jsou dva úseky **a** a **-b** jejichž integrál je nulový. Např. pro $n=3$ se bude nejvyšší derivace modulační funkce skládat z osmi konstantních úseků: $\{a \ -b \ | \ -a \ b \ || \ -a \ b \ a \ -b\}$ (viz. obrázek III.3-2 nahoře). Nižší derivace následovně získáme postupnou numerickou integrací. Modulační funkce pro soustavu třetího řádu jsou znázorněny v následujícím grafu (viz. obrázek III.3-2).



obrázek III.3-2: Modulační funkce a její derivace (ze zdola: $\phi(t)$, $\phi'(t)$, $\phi''(t)$, $\phi'''(t)$)

Ilustrační Příklad:

Vyzkoušejme derivační metodu na stejném příkladě jako u integrační metody: mějme naměřenou přechodovou charakteristiku (známo: $u(t_i)$, $y(t_i)$) soustavy druhého řádu popsané diferenciální rovnicí ve tvaru:

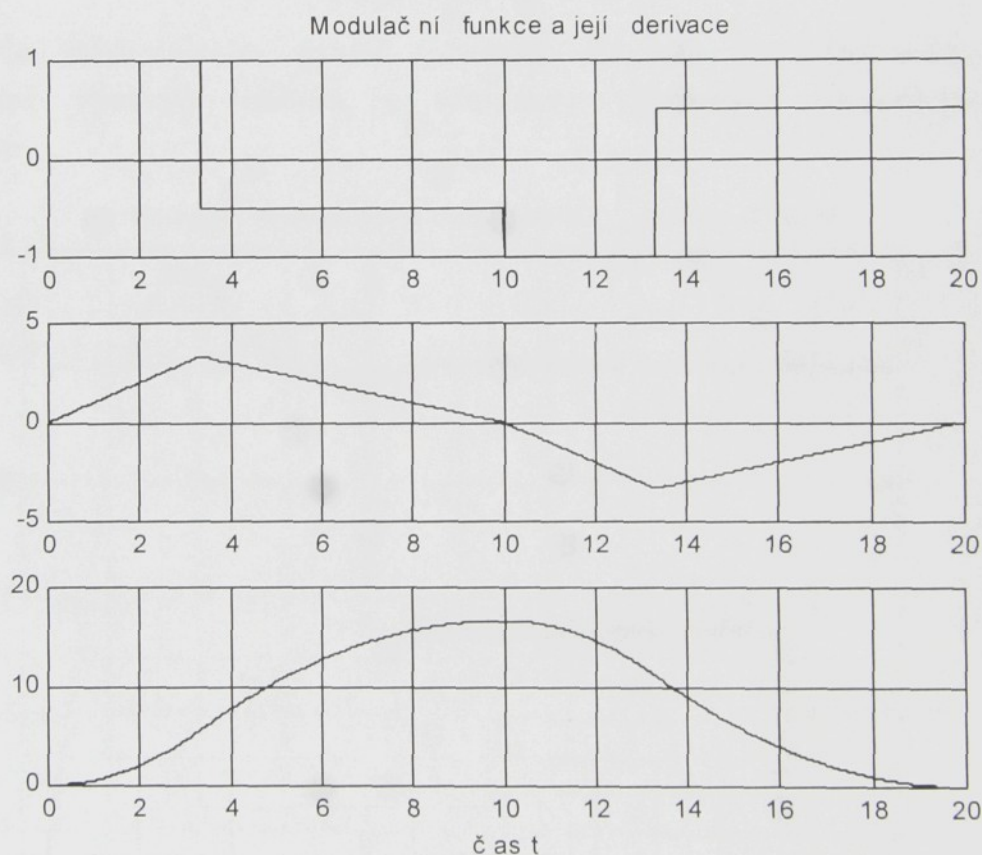
$$5u = \ddot{y} + 1\dot{y} + 1y,$$

kterou tu můžeme přepsat následovně:

$$y = b_0 u - a_1 \dot{y} - a_2 \ddot{y},$$

po integraci a vynásobení modulační funkcí dostaneme vztah, který ještě stále obsahuje derivace měřených veličin:

$$\int_0^{t_f} y(t)\Phi(t)dt = b_0 \int_0^{t_f} u(t)\Phi(t)dt - a_1 \int_0^{t_f} \dot{y}(t)\Phi(t)dt - a_2 \int_0^{t_f} \ddot{y}(t)\Phi(t)dt,$$



**obrázek III.3-3: Příklad modulační funkce a jejích derivací (ze
zdola: $\phi(t)$, $\phi'(t)$, $\phi''(t)$)**

což můžeme za předpokladu o modulační funkci ($\dot{\Phi}(0) = \ddot{\Phi}(t_f) = 0$) přepsat ve tvaru bez derivací měřených signálů:

$$\int_0^{t_f} y(t)\Phi(t)dt = b_0 \int_0^{t_f} u(t)\Phi(t)dt + a_1 \int_0^{t_f} y(t)\dot{\Phi}(t)dt - a_2 \int_0^{t_f} y(t)\ddot{\Phi}(t)dt$$

a z této rovnice pro n lineárně nezávislých modulačních funkcí (viz. obrázek III.3-3) dostaneme n lineárních rovnic pro odhad vektoru parametrů $\mathbf{p}$. Lineárně nezávislých modulačních funkcí potřebujeme nejméně tolik, kolik je neznámých parametrů, v našem případě nejméně tři. Pro vydatnější odhad parametrů jich ale použijeme několikanásobně více.

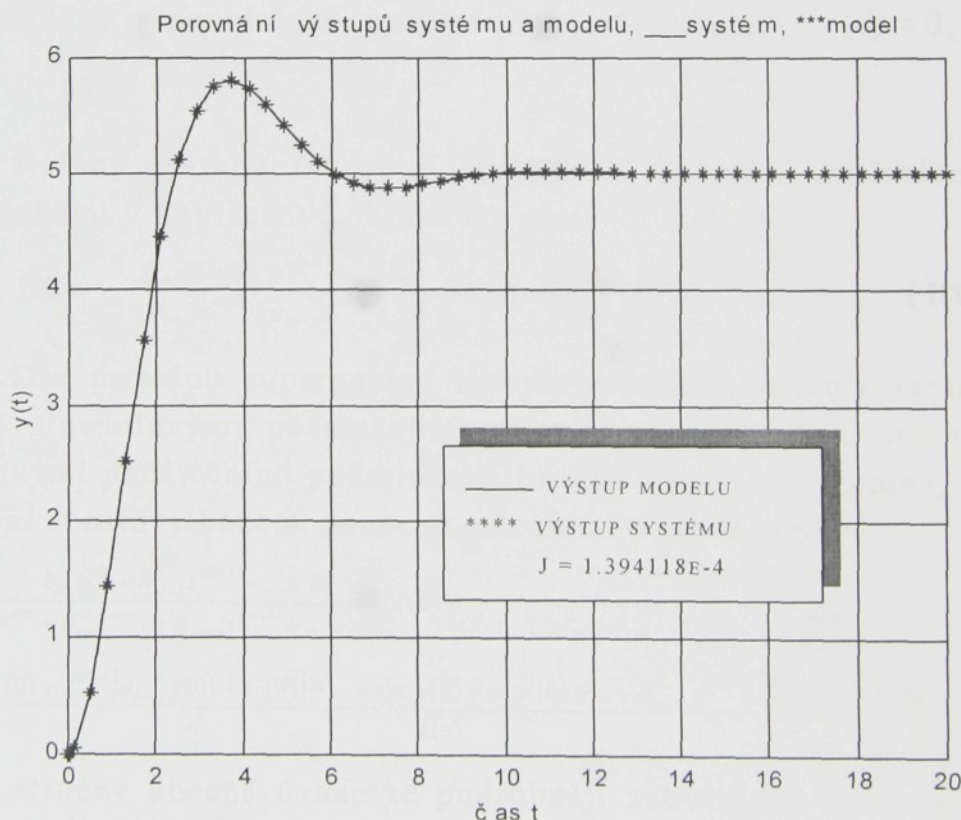
Po provedené optimalizaci metodou nejmenších čtverců jsme našli optimální parametry takto:

$$p(1) = b_0 = 5.0001$$

$$p(2) = a_1 = 1.0016$$

$$p(3) = a_2 = 0.9958.$$

Na následujícím grafu (obrázek III.3-4) je znázorněno porovnání výstupů modelu a systému vybuzených skokovým signálem.



obrázek III.3-4: Porovnání výstupů modelu a systému

III.4 VLIV POČÁTEČNÍCH PODMÍNEK

V praxi se mnohdy setkáváme s technologiemi, u kterých nejsme schopni ovlivnit vstupní signál natolik, abychom mohli započít měření z nulových počátečních podmínek. Situace se dále komplikuje v případě, když je v systému ještě navíc akumulována energie, tzn. že ani vnitřní stavové veličiny nejsou nulové na počátku měření. S nenulovou počáteční podmínkou na výstupu systému jsme schopni se vcelku snadno vyrovnat (např. vyloučením tzv. driftu [OLE]). Horší je tomu s nenulovými počátečními podmínkami vyšších řádů (např. u pohybujícího se tělesa s rychlostí a zrychlením). Při dostatečně dlouhém souboru dat se dají zanedbat, zvláště když použijeme iterační metodu s exponenciálním zapomínáním, ale i tak ovlivňují odhad parametrů soustavy. V této kapitole bude předveden postup, jak se lze s tímto problémem vyrovnat.

Uvažujme znovu lineární systém popsany diferenciální rovnicí ve tvaru (III-20):

$$y^{(na)}(t) + \sum_{i=0}^{na-1} a_i y^{(i)}(t) = \sum_{i=0}^{nb} b_i u^{(i)}(t), \quad (\text{III-20})$$

ale tentokrát s nenulovými počátečními podmínkami $y^{(i)}(0) \neq 0$, $i = 0, 1, \dots, na-1$.

Řešení výstupu takového systému s nulovými počátečními podmínkami v Laplaceově obraze lze psát:

$$Y(s) = \frac{b_{nb}s^{nb} + b_{nb-1}s^{nb-1} + \dots + b_0}{s^{na} + a_{na-1}s^{na-1} + a_{na-2}s^{na-2} + \dots + a_0} U(s) = \frac{B(s)}{A(s)} U(s). \quad (\text{III-21})$$

Dle principu superpozice lze pro lineární systémy rozložit řešení s nenulovými počátečními podmínkami na řešení soustavy s nulovými počátečními podmínkami buzené vstupním signálem $u(t)$ a řešení soustavy buzené pouze počátečními podmínkami:

$$Y(s) = \frac{b_{nb}s^{nb} + b_{nb-1}s^{nb-1} + \dots + b_0}{s^{na} + a_{na-1}s^{na-1} + a_{na-2}s^{na-2} + \dots + a_0} U(s) + \frac{s^{na} y(0) + s^{na-1} [a_{na-1} y(0) + y'(0)] + \dots + s [a_1 y(0) + a_2 y'(0) + \dots + a_{na-1} y^{(na-2)}(0) + y^{(na-1)}(0)]}{A(s)} \frac{1}{s} A.$$

Po stručné obecné úvaze se podrobněji zabývejme tím, jak se nenulové počáteční podmínky projeví v integračním a derivačním identifikačním algoritmu (viz kapitoly III.4.1 a III.4.2).

III.4.1 Aplikace nenulových počátečních podmínek na integrační metodu

Přepíšeme-li řešení jednorozměrného systému s nenulovými počátečními podmínkami a upravíme-li ho pro integrační metodu dostaneme (III-22):

$$y(t) = \sum_{i=0}^{nb} b_i \int_{na-i} u(t) dt - \sum_{i=1}^{na} a_i \int_{na-i+1} y(t) dt + \sum_{i=0}^{na-1} k_i \int_i \eta(t) dt, \quad (\text{III-22})$$

kde

$$k_0 = y(0)$$

$$k_1 = a_{na-1} y(0) + y'(0)$$

$$k_2 = a_{na-2} y(0) + a_{na-1} y'(0) + y''(0)$$

$$k_3 = a_{na-3} y(0) + a_{na-2} y'(0) + a_{na-1} y''(0) + y'''(0)$$

:

$$k_{na-1} = a_1 y(0) + a_2 y'(0) + \dots + a_{na-1} y^{(na-2)}(0) + y^{(na-1)}(0),$$

což můžeme znovu přepsat v maticovém tvaru:

$$\mathbf{y} = \mathbf{X} \mathbf{p}, \quad (\text{III-23})$$

$$\mathbf{X} = \begin{bmatrix} \int_{na} u(t) dt, & \int_{na-1} u(t) dt, \dots, & \int_{na-nb} u(t) dt & | & - \int_{na} y(t) dt, - \int_{na-1} y(t) dt, \dots, \\ & & & & - \int_{na-1} y(t) dt & | & \int \eta(t), \dots, \int_{na-1} \eta(t) \end{bmatrix},$$

$$\mathbf{p} = [b_0, b_1, \dots, b_{nb} \mid a_1, a_2, \dots, a_{na} \mid k_0, k_1, \dots, k_{na-1}]^T.$$

Zahrnutím nenulových počátečních podmínek do struktury modelu se vektor hledaných parametrů rozšířil o na koeficientů k_i . Z těchto koeficientů jsme případně schopni vypočítat počáteční podmínky. Ty nás ale ve většině případů nezajímají. Zvětšení počtu parametrů u metody nejmenších čtverců, pro kterou známe analytickou formuli pro optimální parametry, nehraje z numerického tak velkou roli a neklade přehnané nároky na dobu výpočtu. Jinak je tomu u obecných optimalizačních metod, kde každé zvýšení dimenze prohledávaného prostoru značně ztěžuje hledání optima účelové funkce [MAN].

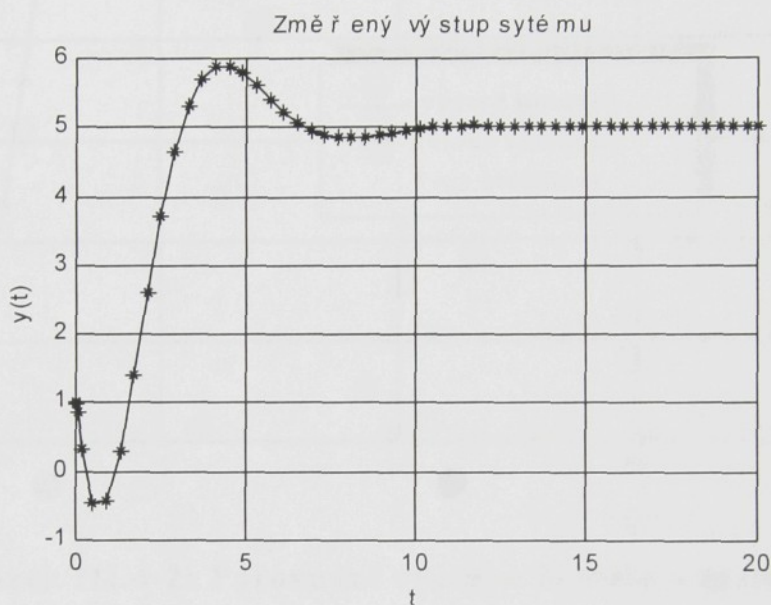
Ilustrační Příklad:

Jako obvykle ilustrujeme použití rozšířeného algoritmu na jednoduchém příkladě: mějme naměřenou (simulovanou) přechodovou charakteristiku (známo: $u(t_i)$, $y(t_i)$) soustavy druhého řádu popsané diferenciální rovnicí ve tvaru:

$$5u = \ddot{y} + 1\dot{y} + 1y,$$

s počátečními podmínkami: $y(0)=1$, $y'(0)=-5$.

Reakce systému na jednotkový skok je znázorněna viz. obrázek III.4-1:



obrázek III.4-1: Změřený výstup

Model uvažujme ve tvaru:

$$y_m = Xp,$$

kde

$$X = \left[\int_0^T u^2 dt, -\int_0^T y^2 dt, -\int_0^T y \dot{y} dt, \int_0^T \dot{y}^2 dt \right], \quad p = [b_0, a_0, a_1, k_0, k_1]^T.$$

Po provedené optimalizaci metodou nejmenších čtverců jsme našli optimální parametry takto:

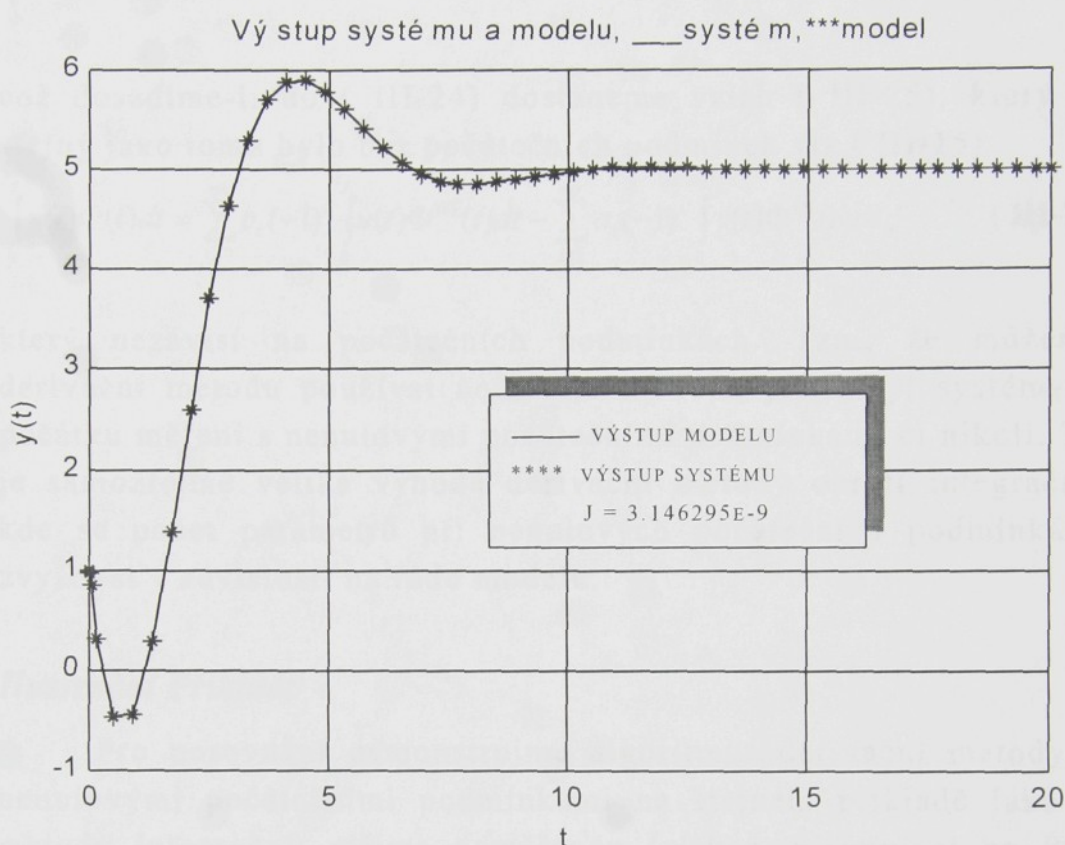
$$p(1) = b_0 = 5.000041, \quad p(2) = a_0 = 1.000008,$$

$$p(3) = a_1 = 1.000016,$$

$$p(4) = k_0 = 1.000000, \quad p(5) = k_1 = -3.999950.$$

Navíc lze lehce spočítat, že koeficienty k_0 , k_1 odpovídají přibližně počátečním podmínkám $y(0)=1$, $y'(0)=-5$.

Porovnáme-li výstup modelu a systému při stejných počátečních podmínkách a buzení dostaneme téměř identické charakteristiky zobrazené na následujícím grafu (obrázek III.4-2):



obrázek III.4-2: Porovnání výstupů modelu a systému

III.4.2 Aplikace nenulových počátečních podmínek na derivační metodu

Přepíšeme-li řešení jednorozměrného systému s nenulovými počátečními podmínkami a upravíme-li ho pro derivační metodu dostaneme (III-24):

$$\int_0^{t_f} y(t) \Phi(t) dt = \sum_{i=0}^{nb} b_i (-1)^i \int_0^{t_f} u(t) \Phi^{(i)}(t) dt - \sum_{i=1}^{na} a_i (-1)^i \int_0^{t_f} y(t) \Phi^{(i)}(t) dt + \sum_{i=0}^{na-1} k_i \int_0^{t_f} \eta(t) \Phi^{(na-i)}(t) dt. \quad (\text{III-24})$$

A protože testovací (modulační) funkce byly definovány tak, že platí:

$$\int_0^{t_f} \Phi^{(i)}(t) dt = 0 \quad \text{pro } i = 1, 2, \dots, na, \text{ pak i: } k_i \int_0^{t_f} \eta(t) \Phi^{(i)}(t) dt = 0,$$

což dosadíme-li do (III-24) dostaneme vztah (III-25), který je stejný jako tomu bylo bez počátečních podmínek viz. (III-15).

$$\int_0^{t_f} y(t) \Phi(t) dt = \sum_{i=0}^{nb} b_i (-1)^i \int_0^{t_f} u(t) \Phi^{(i)}(t) dt - \sum_{i=1}^{na} a_i (-1)^i \int_0^{t_f} y(t) \Phi^{(i)}(t) dt, \quad (\text{III-25})$$

který nezávisí na počátečních podmínkách. Tzn., že můžeme derivační metodu používat nezávisle na tom, jestli byl systém na počátku měření s nenulovými počátečními podmínkami či nikoli. To je samozřejmě velká výhoda derivační metody oproti integrační, kde se počet parametrů při nenulových počátečních podmínkách zvyšoval v závislosti na řádu modelu.

Ilustrační Příklad:

Pro porovnání demonstrujme algoritmus derivační metody s nenulovými počátečními podmínkami na stejném příkladě jako u metody integrační: mějme naměřenou (získanou simulací na PC) přechodovou charakteristiku (známo: $u(t_i)$, $y(t_i)$) soustavy druhého řádu popsané diferenciální rovnicí ve tvaru:

$$5u = \ddot{y} + 1\dot{y} + 1y$$

s počátečními podmínkami: $y(0)=1$, $y'(0)=-5$.

Model uvažujeme v následujícím tvaru (počáteční podmínky vzhledem k vlastnostem modulační funkce neovlivní výpočet optimálních parametrů, tzn. že postupujeme stejně jako u ilustračního příkladu v kapitole III.3):

$$y = b_0 u - a_1 \dot{y} - a_2 \ddot{y},$$

po integraci a vynásobení modulační funkcí dostaneme:

$$\int_0^{t_f} y(t) \Phi(t) dt = b_0 \int_0^{t_f} u(t) \Phi(t) dt - a_1 \int_0^{t_f} \dot{y}(t) \Phi(t) dt - a_2 \int_0^{t_f} \ddot{y}(t) \Phi(t) dt,$$

což můžeme za předpokladu o modulační funkci ($\dot{\Phi}(0) = \ddot{\Phi}(t_f) = 0$) přepsat následovně:

$$\int_0^{t_f} y(t)\Phi(t)dt = b_0 \int_0^{t_f} u(t)\Phi(t)dt + a_1 \int_0^{t_f} y(t)\dot{\Phi}(t)dt - a_2 \int_0^{t_f} y(t)\ddot{\Phi}(t)dt$$

a z této rovnice pro n lineárně nezávislých modulačních funkcí dostaneme n lineárních rovnic pro odhad vektoru parametrů $\mathbf{p}$. Tyto rovnice zpracujeme standardním postupem metody nejmenších čtverců.

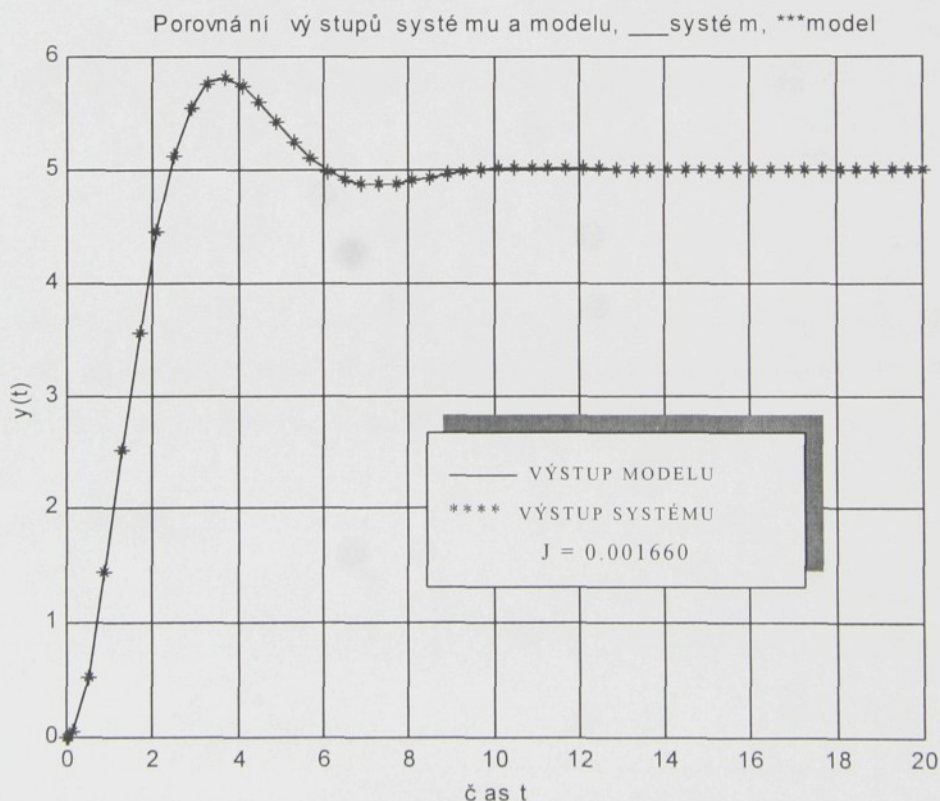
Po provedené optimalizaci metodou nejmenších čtverců byly nalezeny optimální parametry takto:

$$p(1) = b_0 = 5.0001,$$

$$p(2) = a_1 = 1.0012,$$

$$p(3) = a_2 = 0.9982.$$

Bohužel nám tato metoda neumožní výpočet počátečních podmínek, takže ani nemůžeme srovnat změřený výstup systému s modelem. V tomto případě, ale známe tzv. "pravdivý systém", tzn. že můžeme porovnat jeho výstup s výstupem modelu (viz. obrázek III.4-3). A jak již napovídaly nalezené parametry, reakce modelu a systému na stejný budicí signál jsou téměř identické.



obrázek III.4-3: Porovnání výstupů modelu a systému

III.5 INTEGRAČNÍ A DERIVAČNÍ METODA PRO NELINEÁRNÍ SYSTÉMY

Nasnadě je jednoduché rozšíření výše popsaných identifikačních metod na nelineární modely, které ale jsou lineární v parametrech. Pro takto definované struktury modelů lze použít metodu nejmenších čtverců společně s derivační popř. integrační metodou. Vyjádřeme jednorozměrný model ve tvaru (III-26):

$$\frac{d^{na}}{dt^{na}} y(t) + \sum_{i=0}^{na-1} a_i \frac{d^i}{dt^i} y(t) = \sum_{i=0}^{nb} b_i \frac{d^i}{dt^i} u(t) + \sum_{i=1}^{nm} p_i z_i(t, u, y), \quad (\text{III-26})$$

kde $z_i(t, u, y)$ jsou nelineární funkce, o kterých předpokládáme, že na nich model závisí lineárně (přes hledané koeficienty). Příkladem takové nelineární funkce může být např. kvadrát ze vstupního signálu u^2 , odmocnina z tohoto signálu $u^{1/2}$, součin vstupního a výstupního signálu $y \cdot u^2$, atd.

Pro *integrační metodu* definujme výstup modelu v maticovém tvaru (III-27):

$$\mathbf{y}_m = \mathbf{X} \mathbf{p}, \quad (\text{III-27})$$

kde

$$\mathbf{X} = \begin{bmatrix} \int_{na} \mathbf{u}(t_i) dt, & \int_{na-1} \mathbf{u}(t_i) dt, \dots, \int_{na} \mathbf{u}(t_i) dt, & - \int_{na} \mathbf{y}(t_i) dt, & - \int_{na-1} \mathbf{y}(t_i) dt, \dots, & - \int_{na} \mathbf{y}(t_i) dt \\ \int_{na} \mathbf{z}_1(t_i) dt, \dots, & \int_{na} \mathbf{z}_{nm}(t_i) dt \end{bmatrix},$$

$$\mathbf{p} = [b_0, b_1, \dots, b_{nb} \mid a_0, a_1, \dots, a_{na-1} \mid p_1, \dots, p_{nm}]^T.$$

Podobně můžeme sestavit model pro derivační metodu. Takto sestavené modely nám umožní použít metodu nejmenších čtverců a následné nalezení vektoru optimálních parametrů.

Ilustrační Příklad:

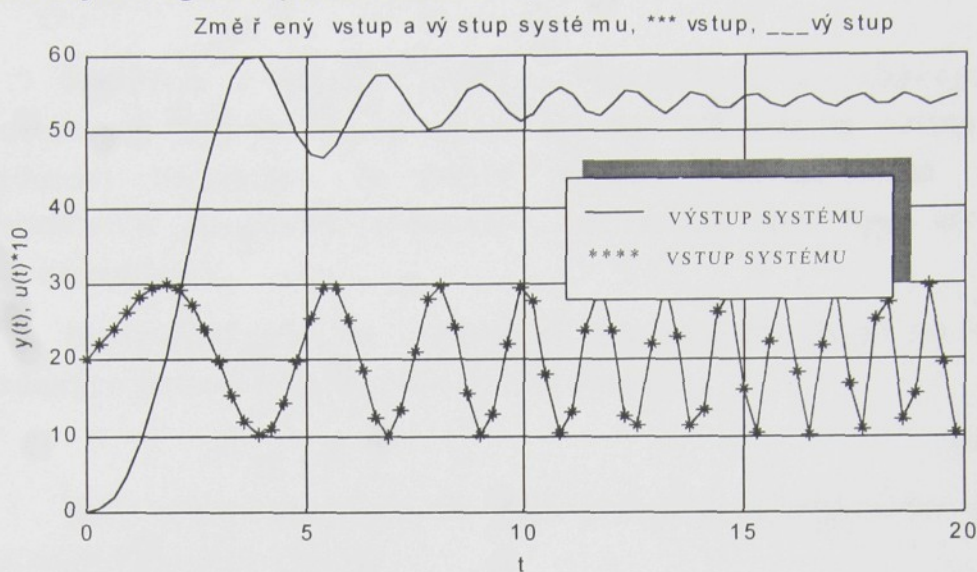
Mějme naměřenou (získanou simulací) reakci nelineárního dynamického systému 2.řádu popsaného následující diferenciální rovnicí:

$$\ddot{y} + a_1 \dot{y} + a_0 y = b_0 u + p \sqrt{y} u,$$

kde je vektor hledaných parametrů dán následovně:

$$\mathbf{p} = [b_0, a_0, a_1, p]^T = [5 \quad 1 \quad 1 \quad 3].$$

Naměřená vstupně-výstupní charakteristika je zobrazena na následujícím grafu (obrázek III.5-1):



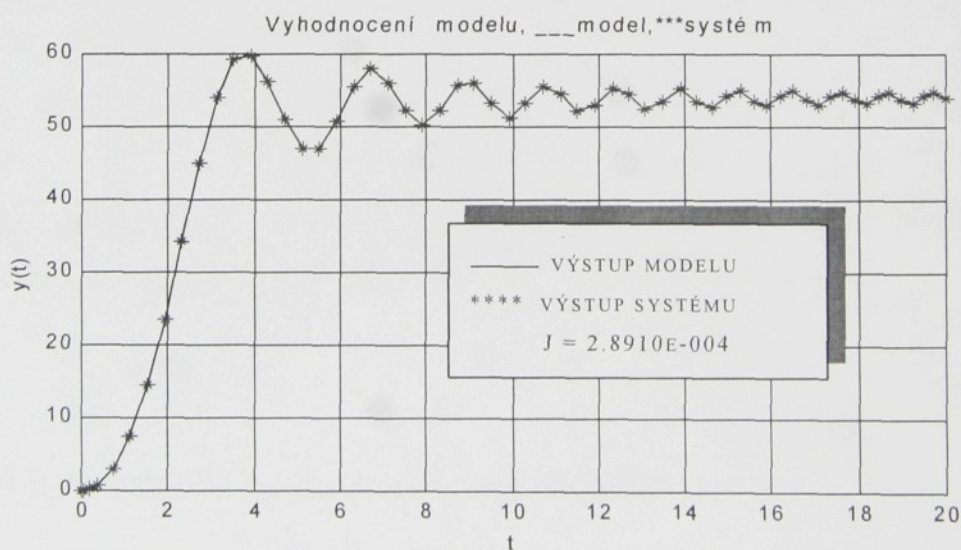
obrázek III.5-1: Změřená vstupně-výstupní charakteristika

Výstup modelu je dán následovně:

$$y_m(t) = b_0 \int_0^t u(t) dt - a_0 \int_0^t y(t) dt - a_1 \int_0^t y(t) dt + p \int_0^t \int_0^t \sqrt{y(t)} u(t) dt.$$

Po provedené optimalizaci metodou nejmenších čtverců jsme našli hledané parametry takto:

$$\hat{p} = [b_0, a_0, a_1, p]^T = [4.9998 \quad 1.0000 \quad 1.0000 \quad 3.0002]$$



obrázek III.5-2: Porovnání výstupů modelu a systému

Nalezené parametry jsou velice blízké hodnotám parametrů skutečného systému. A i reakce modelu a systému na stejný budicí signál (obrázek III.5-2) jsou téměř shodné.

III.6 METODA KVAZILINEARIZACE

Přejdeme k jinému přístupu identifikace a zabývejme se soustavami, jejichž výstup závisí alespoň na jednom nelineárním hledaném parametru. V tomto případě budeme řešit odhad optimálních parametrů systému z měřených dat tzv. metodou kvazilinearizace.

Předpokládejme, že výstup ze soustavy $y(t)$ je až na vektor hledaných parametrů $\mathbf{p}$ znám a je dán funkcí:

$$\mathbf{y}(t) = \mathbf{f}(\mathbf{p}, t). \quad (\text{III-28})$$

Dále předpokládejme, že jsme pro studovaný systém zvolili nelineární model ve tvaru:

$$\mathbf{y}_m = \mathbf{f}(\mathbf{p}, t). \quad (\text{III-29})$$

Účelovou funkci definujeme jako součet kvadrátů odchylek modelu od systému:

$$J = \mathbf{e}^T \mathbf{e}, \quad (\text{III-30})$$

$$\mathbf{e} = \mathbf{y} - \mathbf{y}_m = \mathbf{y} - \mathbf{f}(\mathbf{p}, t). \quad (\text{III-31})$$

Protože výstup modelu je nelineární funkcí neznámým parametrům, je třeba ho před použitím lineárního estimátoru nejprve linearizovat pomocí Taylorova rozvoje (uvažujeme pouze první dva členy, zbytek zanedbejme):

$$\mathbf{y}_m = \mathbf{f}(\mathbf{p}, t) \approx \mathbf{y}_m + \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \Delta \mathbf{p}. \quad (\text{III-32})$$

Minimalizací účelové funkce dané (III-30) dostaneme podmínku optimálního kroku:

$$\frac{\partial J}{\partial \mathbf{p}} = 2 \frac{\partial \mathbf{e}^T}{\partial \mathbf{p}} \mathbf{e} = 0, \quad (\text{III-33})$$

a dosazením (III-32) do (III-31) dostáváme:

$$\mathbf{e} = \mathbf{y} - \mathbf{y}_m = \mathbf{y} - \left(\mathbf{y}_m + \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \Delta \mathbf{p} \right). \quad (\text{III-34})$$

Po dosazení (III-34) do (III-33) získáme:

$$2 \frac{\partial \mathbf{y}_m^T}{\partial \mathbf{p}} \left[\mathbf{y} - \left(\mathbf{y}_m + \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \Delta \mathbf{p} \right) \right] = 0, \quad (\text{III-35})$$

$$\text{z něhož } \Delta \mathbf{p} = \left[\frac{\partial \mathbf{y}_m^T}{\partial \mathbf{p}} \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \right]^{-1} \frac{\partial \mathbf{y}_m^T}{\partial \mathbf{p}} (\mathbf{y} - \mathbf{y}_m). \quad (\text{III-36})$$

(III-36) použijeme ve vztahu (III-37) pro úpravu vektoru parametrů iteračním výpočtem :

$$\mathbf{p}^{k+1} = \mathbf{p}^k + \left[\frac{\partial \mathbf{y}_m^T}{\partial \mathbf{p}} \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \right]^{-1} \frac{\partial \mathbf{y}_m^T}{\partial \mathbf{p}} (\mathbf{y} - \mathbf{y}_m^k). \quad (\text{III-37})$$

V případě, že se při iteračním výpočtu vektor parametrů již dále nemění, výpočet ukončíme (dostáváme tzv. bodový odhad parametrů). Nevýhodou této metody je nutnost počátečního odhadu vektoru $\mathbf{p}$ a také to, že konvergence není zaručena. Podmínky konvergence lze vymezit analyticky [SOU].

Ilustrační Příklad:

Mějme naměřenou přechodovou charakteristiku dynamického systému 1.řádu popsaného lineární diferenciální rovnicí:

$$\dot{y} + 1y = 5u$$

nebo jinak:

$$\dot{y} + a_0 y = b_0 u.$$

Průběh přechodové charakteristiky tohoto systému lze vyjádřit následovně:

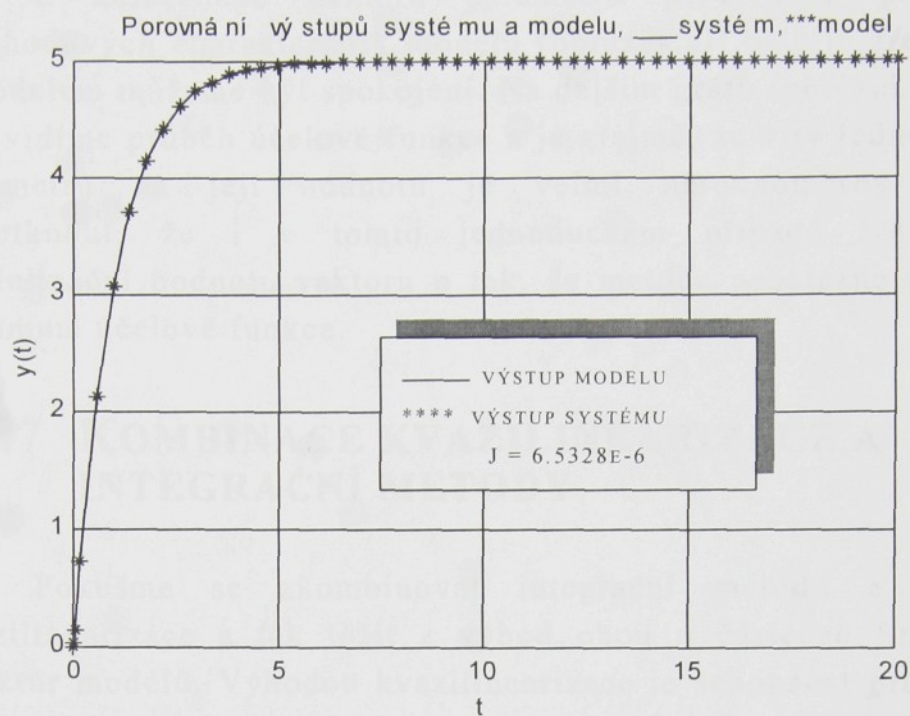
$$y_m = b_0 (1 - e^{-a_0 t}) = p(1) (1 - e^{-p(2)t}),$$

což jsme označili přímo jako výstup modelu. Další členy nutné pro iterační výpočet jsou:

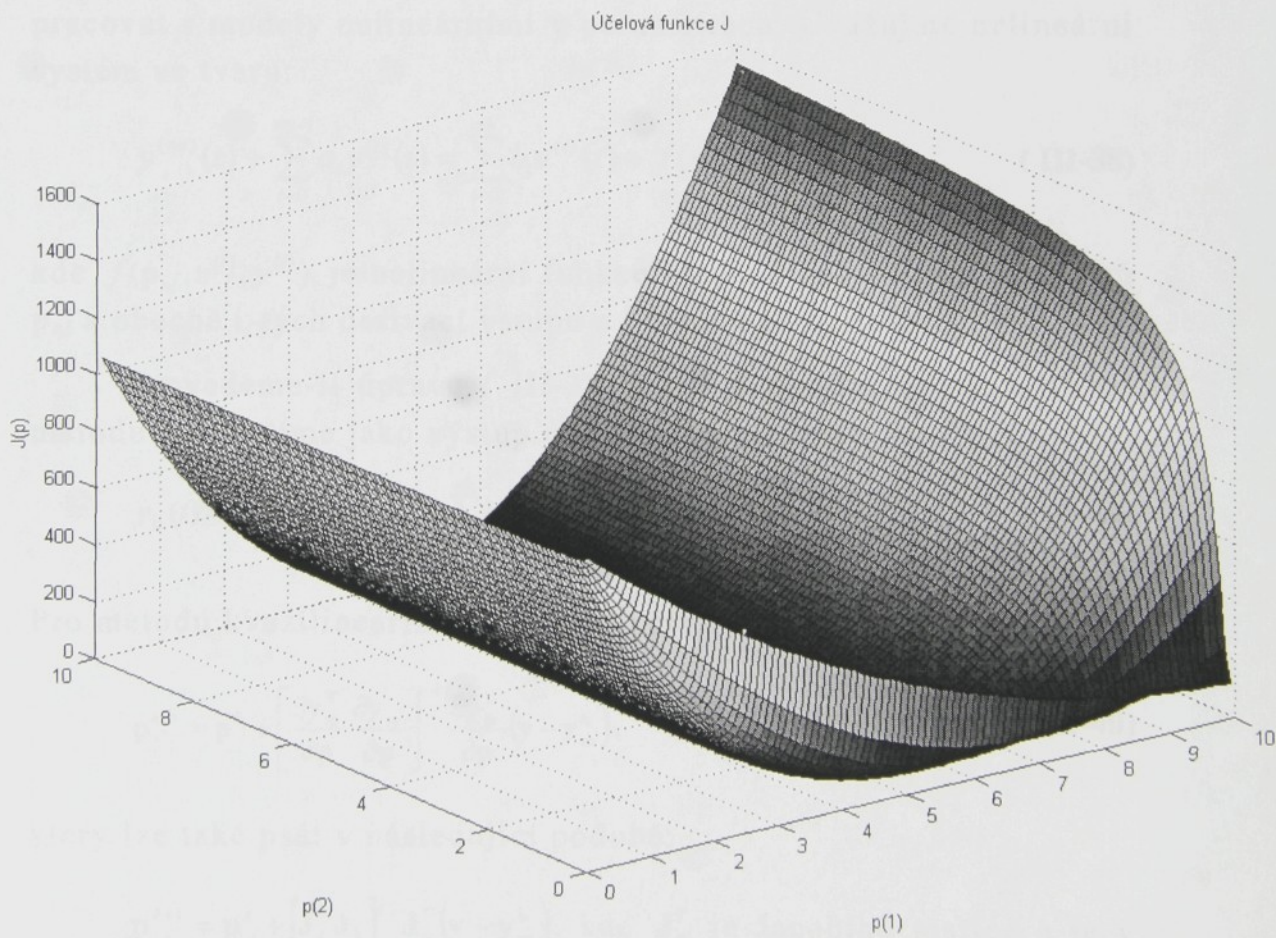
$$\frac{\partial \mathbf{y}_m}{\partial \mathbf{p}}^k = \begin{bmatrix} 1 - e^{-p(2)^k t_1} & p(1)^k t_1 e^{-p(2)^k t_1} \\ \dots & \dots \\ 1 - e^{-p(2)^k t_N} & p(1)^k t_N e^{-p(2)^k t_N} \end{bmatrix}, \quad \mathbf{y}_m^k = \begin{bmatrix} p(1)^k (1 - e^{-p(2)^k t_1}) \\ \dots \\ p(1)^k (1 - e^{-p(2)^k t_N}) \end{bmatrix}$$

Jako počáteční odhad vektoru $\mathbf{p}$ jsme použili hodnoty: $p(1)=1$, $p(2)=0,2$. Již po pěti iteračních krocích jsme získali odhad parametrů velice blízký parametrům studovaného systému:

$$p(1) = b_0 = 5.00000, \quad p(2) = a_0 = 1.00000$$



obrázek III.6-1: Porovnání výstupů modelu a systému



obrázek III.6-2: Průběh účelové funkce

Z nalezeného vektoru parametrů $\mathbf{p}$ a i z porovnání přechodových charakteristik modelu (obrázek III.6-1) je zřejmé, že s modelem můžeme být spokojeni. Na dalším grafu (obrázek III.6-2) pak vidíme průběh účelové funkce a je zřejmé, že vliv jednotlivých parametrů na její hodnotu je velmi nerovnoměrný. Nutno podotknout, že i v tomto jednoduchém případě lze nalézt inicializační hodnotu vektoru $\mathbf{p}$ tak, že metoda nenalezne globální minimum účelové funkce.

III.7 KOMBINACE KVAZILINEARIZACE A INTEGRAČNÍ METODY

Pokusme se zkombinovat integrační metodu a metodu kvazilinearizace a tak těžit z výhod obou u částečně lineárních struktur modelů. Výhodou kvazilinearizace je schopnost pracovat s nelineárními modely, její nevýhodou je vyšší numerická náročnost a nutnost inicializačního odhadu parametrů. Naopak integrační metoda je rychlá a numericky ne příliš náročná, zato ale nedokáže pracovat s modely nelineárními v parametrech. Uvažujme nelineární systém ve tvaru:

$$y^{(na)}(t) + \sum_{i=0}^{na-1} a_i y^{(i)}(t) = \sum_{i=0}^{nb} b_i u^{(i)}(t) + f(\mathbf{p}_{nl}, u^{(i)}, y^{(i)}), \quad (\text{III-38})$$

kde $f(\mathbf{p}_{nl}, u^{(i)}, y^{(i)})$ je nelineární funkce hledaného vektoru parametrů $\mathbf{p}_{nl}$ a obecně i-tých derivací vstupu a výstupu ze systému.

Provedeme-li úpravu (III-38) na tvar vhodný pro integrační metodu a označíme jako výstup modelu, dostaneme:

$$y_m(t) = - \sum_{i=0}^{na-1} a_i \int_{(na-i)} y(t) dt + \sum_{i=0}^{nb} b_i \int_{(na-i)} u(t) dt + \int_{na} f(\mathbf{p}_{nl}, u^{(i)}, y^{(i)}). \quad (\text{III-39})$$

Pro metodu kvazilinearizace jsme v kapitole III.6 získali vztah:

$$\mathbf{p}^{k+1} = \mathbf{p}^k + \left[\frac{\partial \mathbf{y}_m^T}{\partial \mathbf{p}} \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \right]^{-1} \frac{\partial \mathbf{y}_m^T}{\partial \mathbf{p}} (\mathbf{y} - \mathbf{y}_m^k), \quad (\text{III-40})$$

který lze také psát v následující podobě:

$$\mathbf{p}^{k+1} = \mathbf{p}^k + [\mathbf{J}_k^T \mathbf{J}_k]^{-1} \mathbf{J}_k^T (\mathbf{y} - \mathbf{y}_m^k), \text{ kde } \mathbf{J}_k^T \text{ je Jacobiho matice a je v}$$

našem případě dána z (III-39) a definována:

$$\mathbf{J}_k = \left[\int_{na} \mathbf{u} dt, \int_{na-1} \mathbf{u} dt, \dots, \int_{na} \mathbf{y} dt, - \int_{na-1} \mathbf{y} dt, \dots, - \int_{na} \mathbf{y} dt \middle| \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}_{nl}} \right],$$

$$\mathbf{p} = [b_0, b_1, \dots, b_{nb} \mid a_0, a_1, \dots, a_{na-1} \mid \mathbf{p}_{nl}]^T.$$

Možnou aplikaci uvedeného přístupu ilustrujeme na následujícím příkladě.

Ilustrační Příklad:

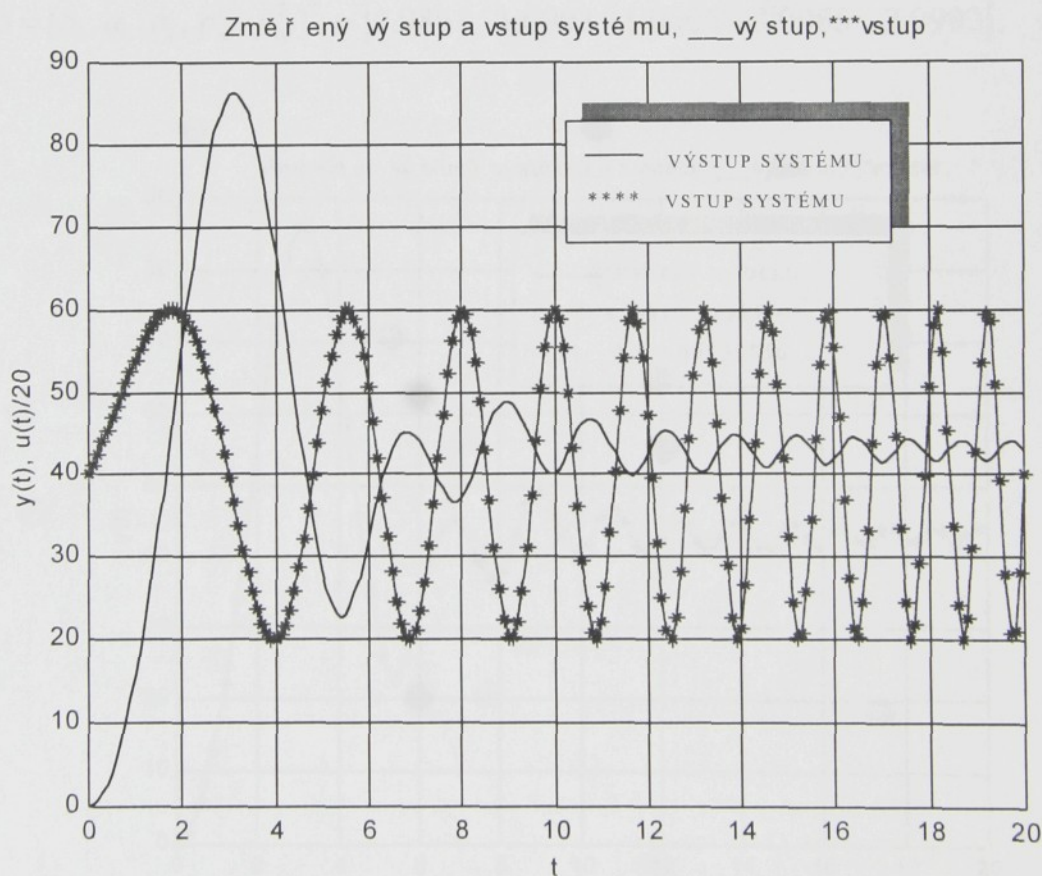
Mějme naměřenou reakci nelineárního dynamického systému 2.řádu popsaného následující diferenciální rovnicí:

$$\ddot{y} + a_1 \dot{y} + a_0 y = b_0 u + p_{nl}^1 u^{p_{nl}^2},$$

kde vektor hledaných parametrů je:

$$\mathbf{p} = [b_0, a_0, a_1, p_{nl}^1, p_{nl}^2]^T = [5 \quad 1 \quad 1 \quad 3 \quad 3].$$

Naměřená vstupně-výstupní charakteristika je zobrazena na následujícím grafu (obrázek III.7-1):



obrázek III.7-1: Změřená vstupně-výstupní charakteristika

Jacobiho matici sestavíme následovně:

$$\mathbf{J}_k = \begin{bmatrix} \int_0^{T_1} \int_0^{T_1} u(t) dt & - \int_0^{T_1} \int_0^{T_1} y(t) dt & - \int_0^{T_1} y(t) dt & \int_0^{T_1} \int_0^{T_1} u(t)^{p_{nl}^2} dt & p_{nl}^1 \int_0^{T_1} \int_0^{T_1} \ln u(T_1) u(t)^{p_{nl}^2} dt \\ \dots & \dots & \dots & \dots & \dots \\ \int_0^{T_N} \int_0^{T_N} u(t) dt & - \int_0^{T_N} \int_0^{T_N} y(t) dt & - \int_0^{T_N} y(t) dt & \int_0^{T_N} \int_0^{T_N} u(t)^{p_{nl}^2} dt & p_{nl}^1 \int_0^{T_N} \int_0^{T_N} \ln u(T_N) u(t)^{p_{nl}^2} dt \end{bmatrix}$$

a výstup modelu je dán následovně:

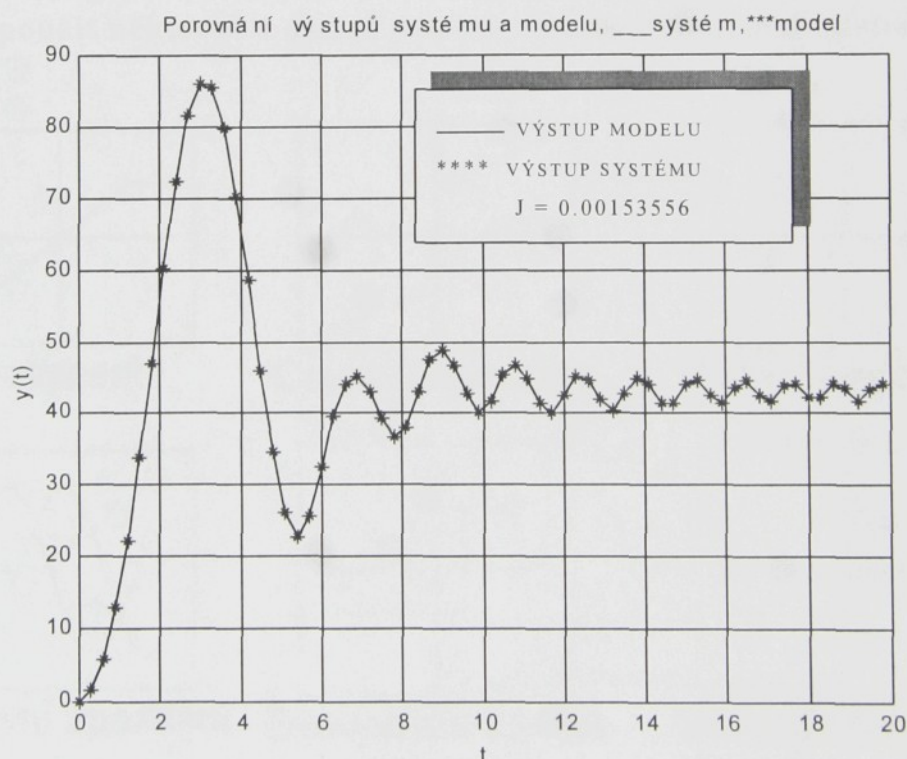
$$\mathbf{y}_m(t) = b_0 \int_0^t \int_0^t u(t) dt - a_0 \int_0^t \int_0^t y(t) dt - a_1 \int_0^t y(t) dt + p_{nl}^1 \int_0^t \int_0^t u(t)^{p_{nl}^2} dt.$$

Prohledávání bylo inicializováno vektorem:

$$\mathbf{p} = [b_0, a_0, a_1, p_{nl}^1, p_{nl}^2]^T = [2 \quad 0.5 \quad 0.5 \quad 2.5 \quad 2.5].$$

Již po čtyřech iteracích jsme získali optimalizované parametry blízké nominálním :

$$\hat{\mathbf{p}} = [b_0, a_0, a_1, p_{nl}^1, p_{nl}^2]^T = [4.9861 \quad 1.0003 \quad 1.0006 \quad 3.0065 \quad 2.9992].$$



obrázek III.7-2: Porovnání výstupů modelu a systému

Tímto způsobem jsme zjistili parametry nelineárního modelu. Metoda konvergovala velice rychle a i výsledný model věrohodně reprodukuje chování studovaného systému (obrázek III.7-2). Ale i v tak jednoduchém ilustračním příkladě narážíme na omezení vyplývající z Jacobiho matice, která obsahuje člen $\ln(u)$. Z toho je jasné, že jsme omezeni pouze na vstupní signál, pro který je definován přirozený logaritmus, tzn. pouze na kladný vstupní signál ($u(t) > 0$). Ale i přes některá omezení je to metoda, která rychle nalezne parametry i nelineárního modelu.

III.8 IDENTIFIKACE NEDIFERENCOVATELNÝCH NELINEÁRNÍCH MODELŮ

V praxi se často setkáváme s nelineárními systémy, jejichž podstata neumožňuje lineární aproximaci. Mezi tyto tzv. "tvrdé nelinearity", které se často objevují v mechatronických systémech, patří Coulombovo tření, nasycení, pásmo necitlivosti, hystereze atd. (obrázek III.8-1). Tato skupina nelinearit má společnou nepříjemnou vlastnost: nelze analyticky určit jejich derivace dle hledaných parametrů. A z toho vyplývající nepříjemným důsledkem je, že pro nalezení optimálního parametru modelu s tvrdou nelinearitou lze jen zřídka použít některou z dosud uvedených identifikačních metod.



obrázek III.8-1: Typické příklady tvrdých nelinearit

Tento problém je proto nutné řešit obecnými optimalizačními metodami. Bohužel náročnost času výpočtu u obecných optimalizačních metod značně roste s počtem hledaných parametrů. V mnoha případech ale můžeme v modelu nalézt lineární část a tu pak řešit metodou nejmenších čtverců a tak snížit počet parametrů hledaných obecnou optimalizační metodou.

Uvažujme popis nelineárního systému ve tvaru (III-41):

$$y^{(na)}(t) + \sum_{i=0}^{na-1} a_i y^{(i)}(t) = \sum_{i=0}^{nb} b_i u^{(i)}(t) + \sum_{i=1}^{nn} p_i z_i(t, u, y) + f(\mathbf{p}_{nl}, u, y, t), \quad (\text{III-41})$$

kde první tři členy tohoto modelu tvoří lineární subsystém, čtvrtý člen $\sum_{i=1}^{nn} p_i z_i(t, u, y)$ je lineární alespoň v parametrech, i když obsahuje nelineární funkce $z_i(t, u, y)$, a konečně poslední člen $f(\mathbf{p}_{nl}, u, y, t)$ je nelineární funkce hledaného vektoru parametrů $\mathbf{p}_{nl}$ a obecně i vstupu a výstupu ze systému a času.

Vektor lineárních parametrů, které budeme hledat metodou nejmenších čtverců, vytvoříme následovně:

$$\mathbf{p} = [b_0, b_1, \dots, b_{nb} \mid a_0, a_1, \dots, a_{na-1} \mid p_1, \dots, p_{nn}]^T.$$

Zbytek hledaných parametrů, vektor $\mathbf{p}_{nl}$ budeme hledat některou obecnou iterační optimalizační metodou. Pro tento vektor také musíme znát prvotní odhad nutný k inicializaci celého algoritmu.

Budeme-li hledat vektor lineárních parametrů $\mathbf{p}$ např. pomocí integrační metody, je výstup modelu dán:

$$\mathbf{y}_m = \mathbf{X}\mathbf{p} + \int_{na} f(\hat{\mathbf{p}}_{nl}^k, u, y, t) dt, \quad (\text{III-42})$$

kde:

$$\mathbf{X} = \left[\int_{na} u dt, \int_{na-1} u dt, \dots, \int_{na} u dt \mid - \int_{na} y dt, - \int_{na-1} y dt, \dots, - \int_{na} y dt \mid \int_{na} z_1 dt, \dots, \int_{na} z_{nn} dt \right],$$

$$\mathbf{p} = [b_0, b_1, \dots, b_{nb} \mid a_0, a_1, \dots, a_{na-1} \mid p_1, \dots, p_{nn}]^T.$$

Celý algoritmus se skládá z následujících kroků:

- $k=1$, volba inicializačního vektoru nelineárních parametrů $\hat{\mathbf{p}}_{nl}^k$.
- Metodou nejmenších čtverců nalezneme vektor optimálních lineárních parametrů $\hat{\mathbf{p}}^k$ pro vektor $\hat{\mathbf{p}}_{nl}^k$.
- Obecnou optimalizační metodu nalezneme $\hat{\mathbf{p}}_{nl}^{k+1}$. Při hledání dosadíme za ostatní parametry modelu vektor $\hat{\mathbf{p}}^k$.
- Zvýšíme k o jedničku a opakujeme algoritmus od druhého kroku.
- Cyklus opakujeme tak dlouho, dokud se nám mění vektor hledaných parametrů (v určité toleranci), popř. dokud jsme nevyčerpali max. počet iterací.

Parametry nalezené tímto algoritmem je vhodné použít jako inicializační pro některou obecnou optimalizační metodu a tak se pokusit nalézt lepší model. Toto doporučení vyplývá z nectností integrační (resp. derivační) metody zmíněné v kapitole III.10.

Možnou aplikaci uvedeného přístupu ilustrujeme na následujícím příkladě.

Ilustrační Příklad:

Uvažujme klasický příklad nelineárního systému vyskytujícího se často při automatizovaném řízení, tzn. lineární systém druhého řádu s dopravním zpožděním.

Přenos takového systému je:

$$F(s) = \frac{b_0}{s^2 + a_1 s + a_0} e^{-sTd}, \quad (\text{III-43})$$

kde Td značí dopravní zpoždění.

Přepíšeme-li přenos (III-43) do časové oblasti dostaneme:

$$\ddot{y}(t) + a_1 \dot{y}(t) + a_0 y(t) = b_0 u(t - Td).$$

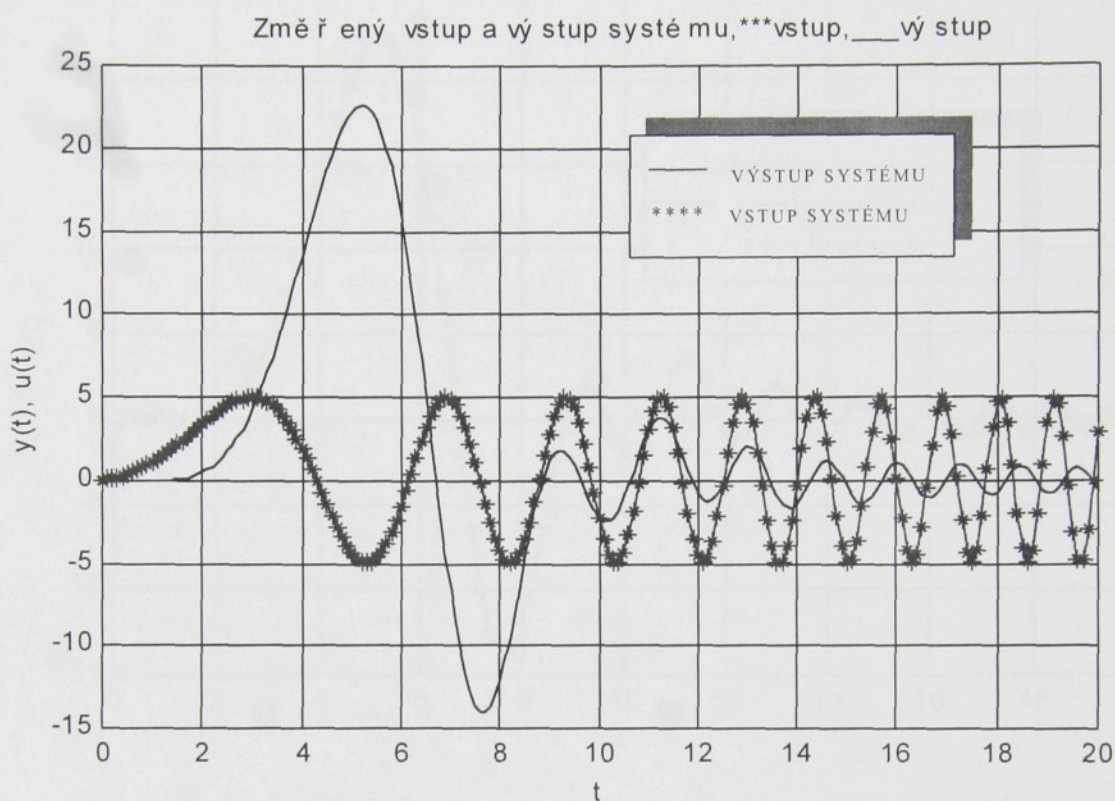
Vektor hledaných lineárních parametrů je tedy dán:

$$\mathbf{p} = [b_0, a_0, a_1]^T = [5 \quad 1 \quad 1]$$

a vektor nelineárních parametrů v tomto případě tvoří pouze jeden prvek:

$$\mathbf{p}_{nl} = [Td] = [1].$$

Naměřená vstupně-výstupní charakteristika je zobrazena na následujícím grafu (obrázek III.8-2):



obrázek III.8-2: Změřená vstupně-výstupní charakteristika

Výstup modelu byl zvolen následovně:

$$y_m(t) = b_0 \int_0^t \int_0^t u(t-Td) dt - a_0 \int_0^t \int_0^t y(t) dt - a_1 \int_0^t y(t) dt.$$

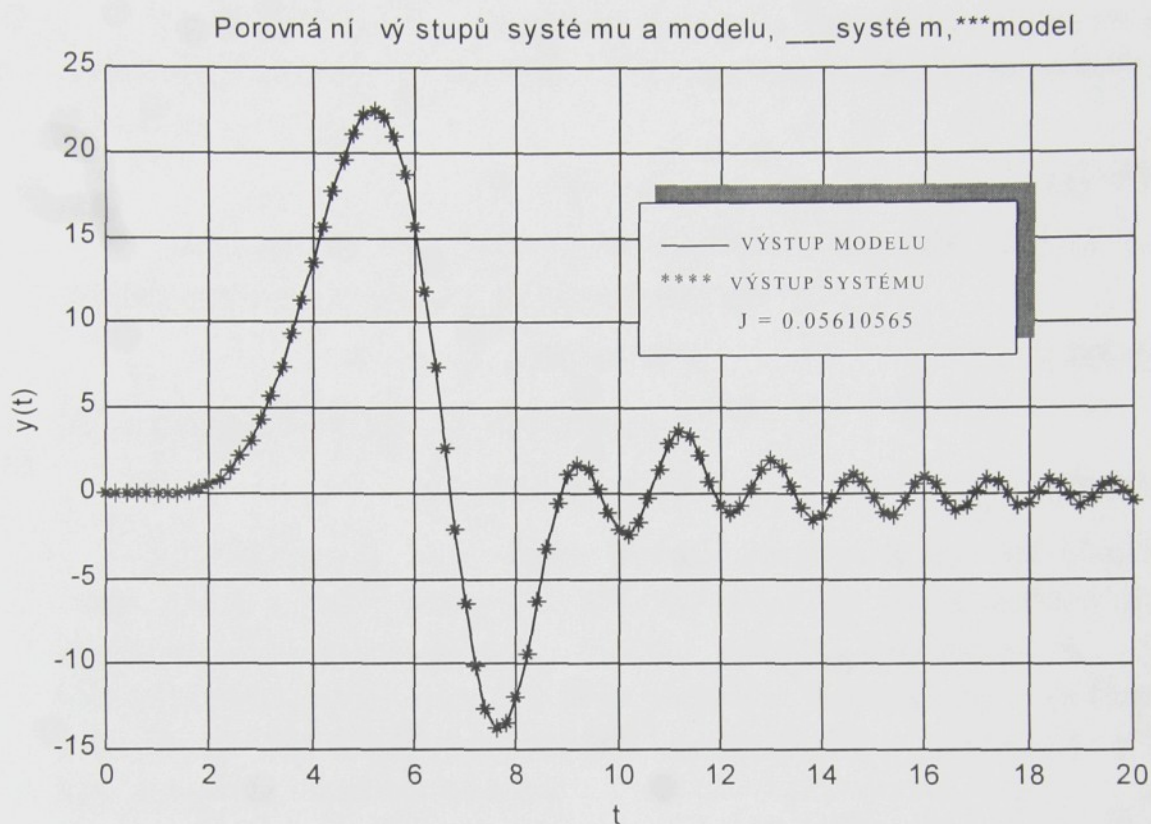
Prohledávání bylo inicializováno vektorem nelineárních parametrů $\hat{\mathbf{p}}_{nl}^1 = [Td] = [0.4]$. Lineární parametry byly následně spočteny metodou nejmenších čtverců pro $\hat{\mathbf{p}}_{nl}^1$. S těmito parametry jsme pomocí Nelder-Mead simplexovou metodou našli optimální parametry $\hat{\mathbf{p}}_{nl}^2$.

Po deseti iteracích jsme získali vektory hledaných parametrů následovně:

$$\mathbf{p} = [b_0, a_0, a_1]^T = [4.9706 \quad 0.9942 \quad 0.9993]$$

$$\mathbf{p}_{nl} = [T] = [0.9970]$$

Na dalším grafu (obrázek III.8-3) je zobrazeno porovnání reakcí modelu a systému na stejný budicí signál.



obrázek III.8-3: Porovnání výstupů modelu a systému

Tímto způsobem jsme zjistili parametry nelineárního modelu. Metoda konvergovala poměrně rychle a i výsledný model věrohodně reprodukuje chování studovaného systému (obrázek III.8-3). Značnou nevýhodou pouze zůstává nutnost inicializovat nelineární parametry.

III.9 ROZBOR Vlivu PORUCH NA ODHAD PARAMETRŮ

Předpokládejme, že naměřená data byla vytvořena systémem s vektorem parametrů $\mathbf{p}^*$ a měření bylo zatíženo náhodnou sekvencí poruch $\mathbf{n}_0$:

$$\mathbf{y} = \mathbf{R}\mathbf{p}^* + \mathbf{n}_0. \quad (\text{III-44})$$

Dle metody nejmenších čtverců (viz. kapitola II.3.2) je estimovaný odhad vektoru parametrů dán dle

$$\hat{\mathbf{p}}_{ls} = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{y}. \quad (\text{III-45})$$

Dosazením (II-44) do (II-45) dostaneme

$$\hat{\mathbf{p}}_{ls} = \mathbf{p}^* + (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{n}_0. \quad (\text{III-46})$$

Požadavkem je, aby vektor odhadu parametrů $\hat{\mathbf{p}}_{ls}$ byl blízký skutečnému $\mathbf{p}^*$ a konvergoval k němu když se počet vzorků na systému blíží k nekonečnu. Vektor odhadu parametrů $\hat{\mathbf{p}}_{ls}$ je nestranným odhadem, tzn. že jeho průměrná hodnota pro všechna data, která by mohla být získána na systému s parametry $\mathbf{p}^*$, konverguje ke skutečné hodnotě

$$E_{y|\mathbf{p}^*} \{\hat{\mathbf{p}}_{ls}\} = \mathbf{p}^*, \quad (\text{III-47})$$

pouze a jen tehdy, jestliže

$$E_{y|\mathbf{p}^*} \{(\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{n}_0\} = \mathbf{0}. \quad (\text{III-48})$$

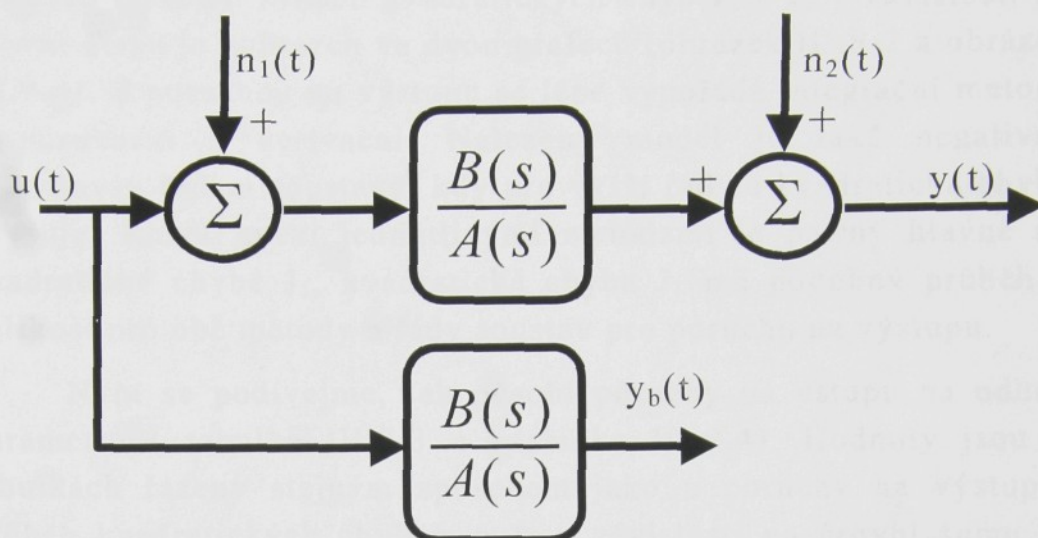
Tzn. proto aby estimovaný odhad parametrů $\hat{\mathbf{p}}_{ls}$ konvergoval ke skutečnému $\mathbf{p}^*$ musí být splněno:

- matice $E\{\mathbf{R}^T \mathbf{R}\}$ není singulární a
- jestliže $E\{\mathbf{R}^T \mathbf{n}_0\} = \mathbf{0}$. Což bude splněno buď, když
 - $\{\mathbf{n}_0\}$ je sekvence nezávislých náhodných poruch s nulovou střední hodnotou (bílý šum). Potom $\{\mathbf{n}_0\}$ nezávisí na tom co se stalo v minulosti a $E\{\mathbf{R}^T \mathbf{n}_0\} = \mathbf{0}$.

nebo

- vstupní signál $\{u(t)\}$ je nezávislý na náhodné poruše $\{\mathbf{n}_0\}$ s nulovou střední hodnotou a $\mathbf{R}$ neobsahuje členy y .

Dále experimentálně rozeberme vliv poruch na odhad parametrů integrační a derivační metodou: uvažujme dvě možná místa, ve kterých mohou poruchy vstupovat do systému (obrázek III.9-1).



obrázek III.9-1: Poruchy vstupující do systému

Poruchu n_1 budeme nazývat poruchou na vstupu, n_2 pak poruchou na výstupu. Výstup systému nezatíženého poruchami, tzv. pravdivý výstup, označme $y_b(t)$. Velikost působící poruchy budeme kvantifikovat tzv. úrovní šumu q , která je definována jako poměr střední hodnoty absolutních hodnot šumu a střední hodnoty absolutních hodnot užitečného signálu vynásobený 100. Vliv poruchy budeme vyhodnocovat pomocí:

- kvadratické chyby mezi modelem a "true" systémem J_1
- kvadratické chyby mezi modelem a měřením J_2

$$J_1(\tau) = \int_0^\tau (y_m(t) - y_b(t))^2 dt, \quad J_2(\tau) = \int_0^\tau (y(t) - y_b(t))^2 dt. \quad (\text{III-49})$$

Pro vyhodnocování vlivu šumu byly použity dvě soustavy:

$\dot{y} + y = 5u$ a $\ddot{y} + \dot{y} + y = 5u$. Analogicky jsme uvažovali modely ve tvaru: $a_1\dot{y} + a_0y = bu$ a $a_2\ddot{y} + a_1\dot{y} + a_0y = bu$ pro integrační metodu, $ad_1\dot{y} + ad_0y = bdu$ a $ad_2\ddot{y} + ad_1\dot{y} + ad_0y = bdu$ pro derivační metodu. Zesílení těchto modelů definujme takto: $k=b/a_0$, $kd=bd/ad_0$.

Nejdříve zhodnoňme vliv poruchy na výstupu na odhad parametrů derivační a integrační metodou pro soustavu 1. řádu (tabulka III.9-1). V této tabulce jsou zapsány hodnoty parametrů

modelů získaných dvěma výše zmíněnými metodami pro různé úrovně šumu. Pro každý model jsou také uvedeny hodnoty kvadratických chyb J_1 a J_2 a zesílení modelu k , respektive kd . Jak ovlivňuje stejná porucha soustavu 2. řádu je uvedeno v další tabulce (tabulka III.9-2). Průběh kvadratických chyb J_1 a J_2 v závislosti na úrovni šumu je zobrazen ve dvou grafech (obrázek III.9-2 a obrázek III.9-3). S poruchou na výstupu se lépe vypořádá integrační metoda v porovnání s derivační. Nalezený model je také negativně ovlivňován řádem soustavy, kdy pro vyšší řád se kvadratická chyba zvyšuje. Rozdíl mezi jednotlivými metodami je patrný hlavně na kvadratické chybě J_1 , kvadratická chyba J_1 má podobný průběh a velikost pro obě metody a řády soustav pro poruchu na výstupu.

Nyní se podívejme, jak působí poruchy na vstupu na odhad parametrů (tabulka III.9-3 a tabulka III.9-4). Hodnoty jsou v tabulkách řazeny stejným způsobem jako u poruchy na výstupu. Průběh kvadratických chyb J_1 a J_2 v závislosti na úrovni šumu je zobrazen ve dvěma grafy (obrázek III.9-4 a obrázek III.9-5). Poruchy na vstupu měly přibližně stejný vliv na kvadratickou chybu J_1 jako poruchy na výstupu. U kvadratické chyby J_2 je tomu již jinak, ta má v tomto případě řádově nižší hodnotu, což lze snadno vysvětlit zatlumením vysokých frekvencí chyb při průchodu soustavou.

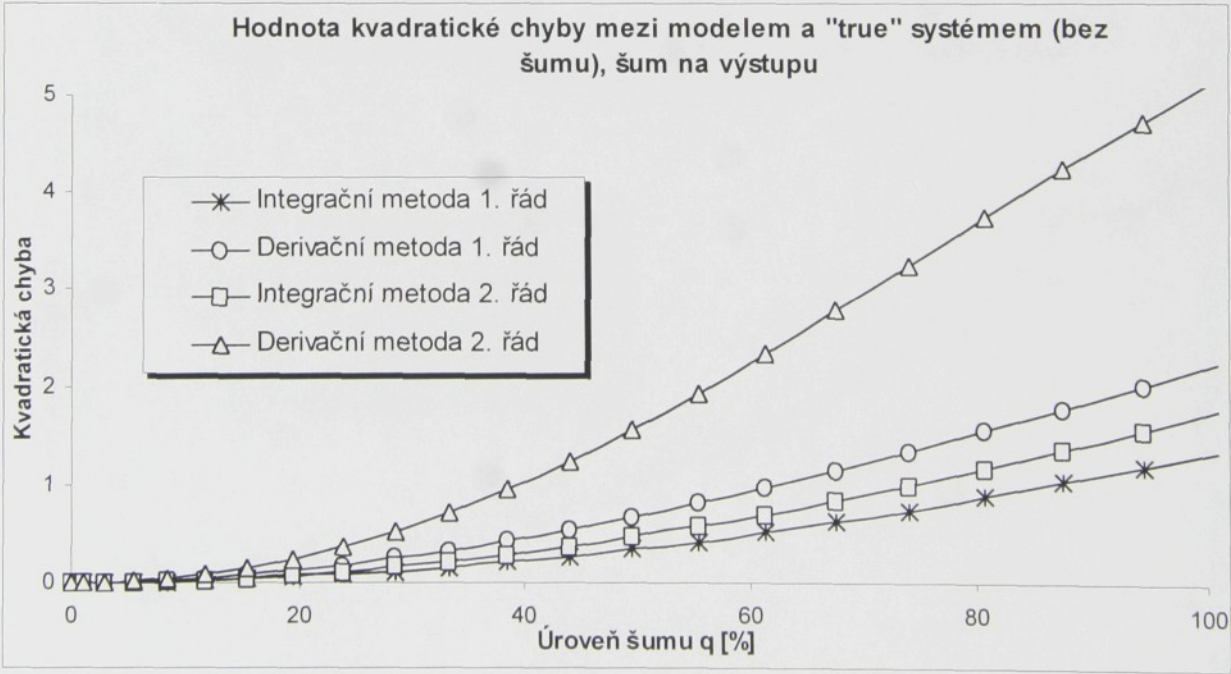
Daleko hůře se nám projeví porucha o nízkých frekvencích (tzv. drift) a to jak na vstupu (obrázek III.9-8 a obrázek III.9-9), tak i na výstupu (obrázek III.9-6 a obrázek III.9-7). Porucha s charakterem driftu na výstupu se projevuje zejména s rostoucím řádem modelu (resp. soustavy). Přibližně stejnou hodnotu kvadratické chyby J_1 dostaneme již pro desetinou úroveň poruchy typu driftu v porovnání s poruchami o vyšších frekvencích (zjednodušeně lze říci, že nízkofrekvenční porucha se nám na kvalitu modelu projevila desetkrát hůře v porovnání s poruchou o vyšších frekvencích). Kvadratická chyba J_1 je u derivační metody pro soustavu 2. řádu (porucha typu driftu) přímo katapultována vzhůru (viz. obrázek III.9-6). Částečným řešením omezení vlivu nízkofrekvenčních poruch je rozšíření modelu o parametry trendu [OLE]. Porucha typu driftu působící na vstupu se projevuje ještě hůře než na výstupu a pro srovnatelnou velikost poruchy kvadratická chyba J_1 vzrůstá na asi pěti násobek. To je v přímé souvislosti se zesílením soustavy, které je v těchto případech také pět.

Posudme ještě vliv periody vzorkování na kvadratickou chybu J_1 na integrační metodu(obrázek III.9-10) a na derivační metodu (obrázek III.9-11). Zatímco u integrační metody je tento vliv poměrně malý a se vzrůstající vzorkovací periodou je monotónní, u derivační metody kvadratická chyba dosahuje značně vyšších hodnot a nemá monotónní průběh. Se vzrůstajícím řádem soustavy kvadratická chyba také vzrůstá. Z tohoto hodnocení vyplývá, že pro stejnou přesnost musíme pro vyšší řády volit nižší vzorkovací periody a to hlavně u derivační metody, což samozřejmě klade vyšší nároky na paměť počítače a dobu výpočtu.

Na závěr shrňme vliv šumu na kvalitu nalezeného modelu integrační a derivační metodou. Zatímco šum o vyšších frekvencích nemá tak výrazný vliv na nalezený model, šum o nižších frekvencích již při menších úrovních působí značně destruktivně na výsledný model. Částečným řešením je rozšíření modelu o parametry trendu. Negativní vliv šumu je dále umocňován řádem identifikované soustavy (respektive modelu). Obecně si lépe vedla integrační metoda v porovnání s derivační a to platí i pro velikost vzorkovací periody, kdy derivační metoda pro stejnou přesnost potřebuje menší vzorkovací periodu.

q	a ₁	a ₀	b	J ₁	J ₂	k	ad ₁	ad ₀	bd	J ₁	J ₂	kd
0,0000	1,0000	1,0000	5,0000	-3,27E-05	-3,27E-05	5,0000	1,0000	1,0011	5,0053	-2,62E-05	-2,62E-05	4,9999
1,0570	1,0000	0,9963	4,9846	1,58E-04	4,82E-02	5,0029	1,0000	0,9929	4,9677	3,22E-04	4,84E-02	5,0030
2,9896	1,0000	0,9897	4,9566	1,26E-03	4,28E-01	5,0081	1,0000	0,9785	4,9009	2,98E-03	4,29E-01	5,0088
5,4923	1,0000	0,9812	4,9208	4,27E-03	1,30E+00	5,0150	1,0000	0,9605	4,8182	1,03E-02	1,30E+00	5,0163
8,4560	1,0000	0,9713	4,8790	1,01E-02	3,22E+00	5,0231	1,0000	0,9403	4,7252	2,42E-02	3,23E+00	5,0252
11,8176	1,0000	0,9603	4,8324	1,98E-02	6,40E+00	5,0323	1,0000	0,9187	4,6257	4,67E-02	6,42E+00	5,0353
15,5346	1,0000	0,9466	4,7734	3,47E-02	1,13E+01	5,0424	1,0000	0,8962	4,5227	7,92E-02	1,13E+01	5,0465
19,5758	1,0000	0,9342	4,7211	5,46E-02	1,65E+01	5,0535	1,0000	0,8734	4,4181	1,23E-01	1,66E+01	5,0586
23,9171	1,0000	0,9213	4,6666	8,11E-02	2,33E+01	5,0654	1,0000	0,8505	4,3135	1,79E-01	2,33E+01	5,0718
28,5389	1,0000	0,9078	4,6101	1,15E-01	4,30E+01	5,0781	1,0000	0,8278	4,2102	2,49E-01	4,31E+01	5,0858
33,4251	1,0000	0,8941	4,5522	1,57E-01	4,82E+01	5,0916	1,0000	0,8056	4,1091	3,33E-01	4,83E+01	5,1006
38,5622	1,0000	0,8801	4,4934	2,08E-01	6,05E+01	5,1058	1,0000	0,7839	4,0108	4,31E-01	6,07E+01	5,1162
43,9385	1,0000	0,8659	4,4338	2,68E-01	8,57E+01	5,1206	1,0000	0,7629	3,9159	5,44E-01	8,58E+01	5,1326
49,5436	1,0000	0,8516	4,3738	3,39E-01	1,05E+02	5,1361	1,0000	0,7427	3,8246	6,73E-01	1,05E+02	5,1497
55,3688	1,0000	0,8372	4,3137	4,22E-01	1,37E+02	5,1523	1,0000	0,7232	3,7371	8,17E-01	1,38E+02	5,1676
61,4059	1,0000	0,8229	4,2537	5,17E-01	1,58E+02	5,1690	1,0000	0,7045	3,6536	9,76E-01	1,59E+02	5,1861
67,6477	1,0000	0,8087	4,1941	6,24E-01	2,13E+02	5,1863	1,0000	0,6866	3,5741	1,15E+00	2,12E+02	5,2053
74,0877	1,0000	0,7945	4,1349	7,45E-01	2,89E+02	5,2042	1,0000	0,6696	3,4986	1,34E+00	2,90E+02	5,2251
80,7201	1,0000	0,7805	4,0764	8,79E-01	2,68E+02	5,2227	1,0000	0,6533	3,4269	1,55E+00	2,69E+02	5,2456
87,5393	1,0000	0,7667	4,0186	1,03E+00	3,87E+02	5,2416	1,0000	0,6378	3,3591	1,77E+00	3,87E+02	5,2666
94,5405	1,0000	0,7530	3,9618	1,19E+00	4,00E+02	5,2611	1,0000	0,6231	3,2949	2,01E+00	3,99E+02	5,2883
101,7190	1,0000	0,7396	3,9059	1,37E+00	4,48E+02	5,2810	1,0000	0,6090	3,2344	2,27E+00	4,48E+02	5,3106

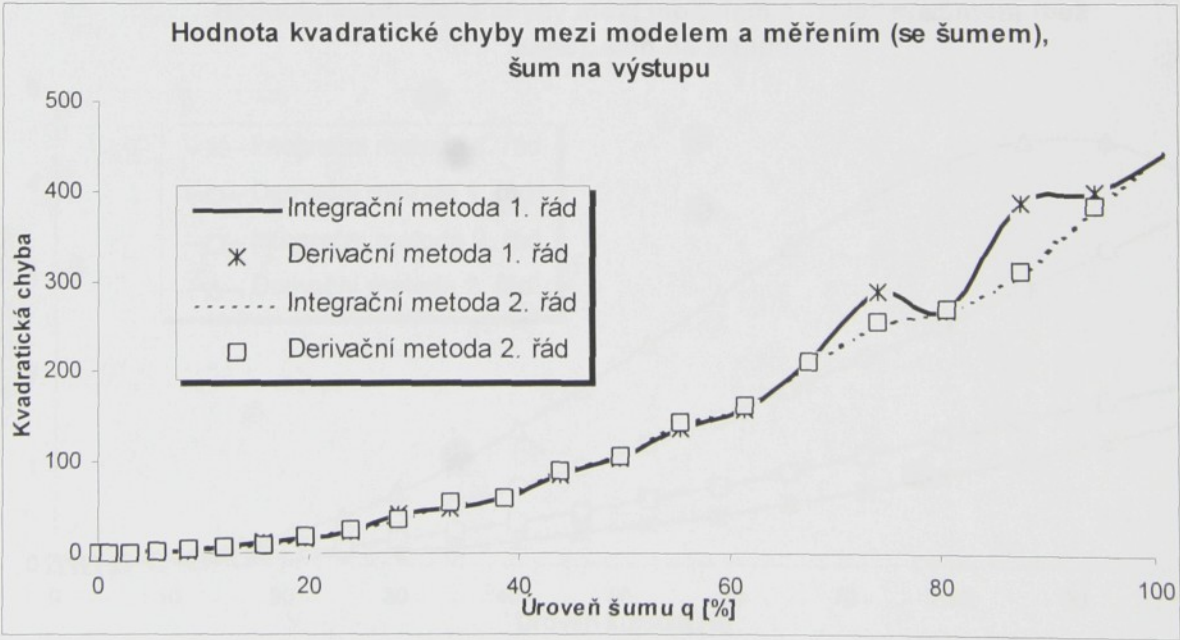
tabulka III.9-1: Vliv poruch na výstupu na derivační a
 integrační metodu pro soustavu 1. řádu



obrázek III.9-2: Kvadratická chyba J₁ v závislosti na úrovni
 poruchy působící na výstupu

q	a ₂	a ₁	a ₀	b	J ₁	J ₂	k	ad ₂	ad ₁	ad ₀	bd	J ₁	J ₂	kd
0,00	1,00	1,00	1,00	5,00	-5,56E-05	-5,56E-05	5,00	1,00	1,00	1,00	4,99	-2,49E-06	-2,49E-06	5,00
1,06	1,00	1,00	1,00	5,01	2,38E-04	4,90E-02	5,00	1,00	1,00	1,00	5,01	5,35E-04	4,96E-02	5,00
2,99	1,00	1,01	1,00	5,02	1,89E-03	4,39E-01	5,01	1,00	1,00	1,01	5,07	5,01E-03	4,41E-01	5,01
5,49	1,00	1,03	1,01	5,04	6,31E-03	1,29E+00	5,02	1,00	1,00	1,02	5,14	1,78E-02	1,31E+00	5,02
8,46	1,00	1,04	1,01	5,06	1,48E-02	3,46E+00	5,02	1,00	1,00	1,04	5,23	4,33E-02	3,51E+00	5,02
11,82	1,00	1,05	1,01	5,09	2,86E-02	6,36E+00	5,03	1,00	1,01	1,06	5,34	8,60E-02	6,42E+00	5,03
15,53	1,00	1,07	1,01	5,11	4,88E-02	1,01E+01	5,04	1,00	1,01	1,08	5,46	1,50E-01	1,01E+01	5,05
19,58	1,00	1,09	1,01	5,13	7,97E-02	1,75E+01	5,06	1,00	1,02	1,11	5,60	2,41E-01	1,78E+01	5,06
23,92	1,00	1,11	1,02	5,15	1,17E-01	2,53E+01	5,07	1,00	1,03	1,14	5,77	3,63E-01	2,56E+01	5,07
28,54	1,00	1,13	1,02	5,18	1,64E-01	3,62E+01	5,08	1,00	1,04	1,17	5,95	5,20E-01	3,69E+01	5,08
33,43	1,00	1,14	1,02	5,20	2,21E-01	5,48E+01	5,09	1,00	1,06	1,21	6,17	7,17E-01	5,56E+01	5,10
38,56	1,00	1,16	1,02	5,22	2,90E-01	5,99E+01	5,11	1,00	1,07	1,25	6,41	9,55E-01	6,04E+01	5,11
43,94	1,00	1,18	1,02	5,24	3,72E-01	9,07E+01	5,12	1,00	1,10	1,30	6,68	1,24E+00	9,23E+01	5,13
49,54	1,00	1,20	1,02	5,25	4,66E-01	1,07E+02	5,14	1,00	1,12	1,36	6,98	1,56E+00	1,08E+02	5,15
55,37	1,00	1,22	1,02	5,26	5,74E-01	1,42E+02	5,16	1,00	1,15	1,42	7,32	1,93E+00	1,45E+02	5,16
61,41	1,00	1,24	1,02	5,27	6,98E-01	1,61E+02	5,17	1,00	1,19	1,49	7,70	2,34E+00	1,63E+02	5,18
67,65	1,00	1,26	1,01	5,26	8,36E-01	2,09E+02	5,19	1,00	1,23	1,56	8,13	2,78E+00	2,12E+02	5,20
74,09	1,00	1,27	1,01	5,26	9,92E-01	2,53E+02	5,21	1,00	1,28	1,65	8,61	3,25E+00	2,58E+02	5,22
80,72	1,00	1,29	1,00	5,25	1,16E+00	2,69E+02	5,23	1,00	1,34	1,74	9,13	3,74E+00	2,72E+02	5,24
87,54	1,00	1,30	1,00	5,23	1,35E+00	3,09E+02	5,25	1,00	1,41	1,85	9,71	4,23E+00	3,13E+02	5,26
94,54	1,00	1,31	0,99	5,20	1,56E+00	3,82E+02	5,27	1,00	1,49	1,96	10,35	4,73E+00	3,86E+02	5,28
101,72	1,00	1,32	0,98	5,17	1,79E+00	4,53E+02	5,29	1,00	1,58	2,08	11,04	5,22E+00	4,57E+02	5,30

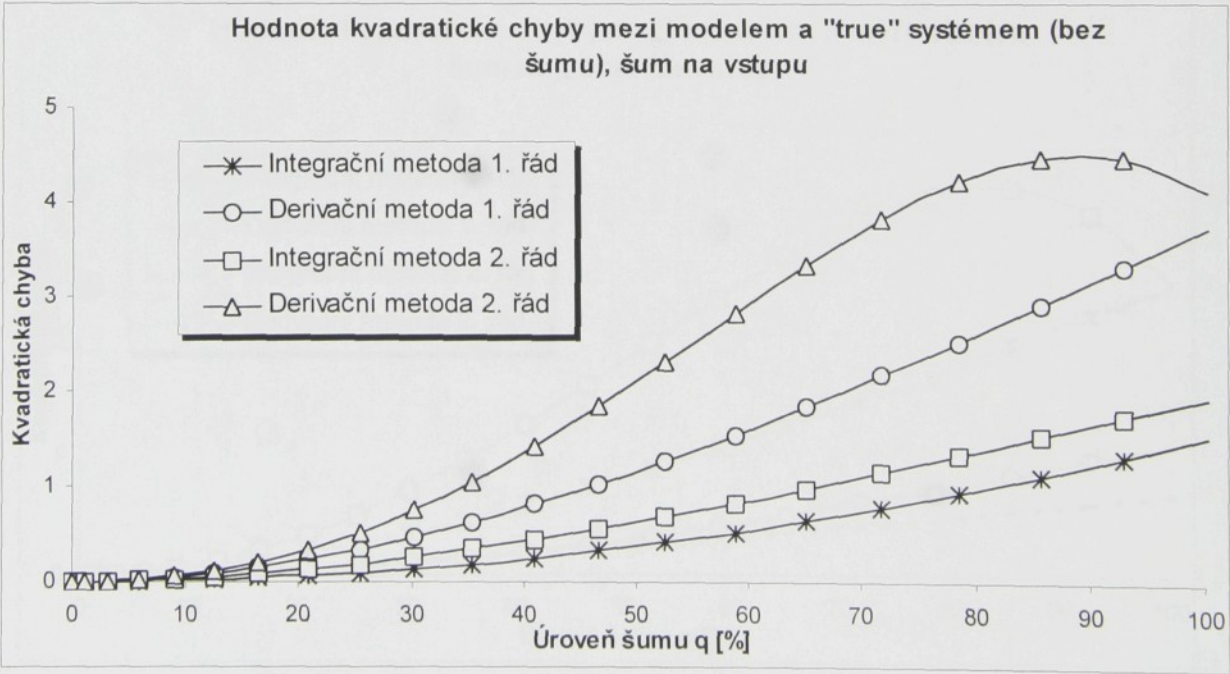
tabulka III.9-2: Vliv poruch na výstupu na derivační a
integrační metodu pro soustavu 2. řádu



obrázek III.9-3: Kvadratická chyba J₂ v závislosti na úrovni
poruchy působící na výstupu

q	a ₁	a ₀	b	J ₁	J ₂	k	ad ₁	ad ₀	bd	J ₁	J ₂	kd
0,00	1,00	1,00	5,00	-3,27E-05	-3,27E-05	5,00	1,00	1,00	5,01	-2,62E-05	-2,62E-05	5,00
1,12	1,00	1,00	4,98	1,62E-04	1,25E-04	5,00	1,00	0,99	4,95	6,22E-04	4,71E-04	5,00
3,18	1,00	0,99	4,95	1,47E-03	1,23E-03	5,01	1,00	0,97	4,85	5,74E-03	4,44E-03	5,01
5,83	1,00	0,98	4,91	4,98E-03	4,41E-03	5,02	1,00	0,94	4,73	1,96E-02	1,55E-02	5,02
8,98	1,00	0,97	4,86	1,18E-02	9,78E-03	5,02	1,00	0,92	4,60	4,62E-02	3,53E-02	5,03
12,55	1,00	0,96	4,81	2,32E-02	2,00E-02	5,03	1,00	0,89	4,46	8,89E-02	6,77E-02	5,04
16,50	1,00	0,94	4,75	4,01E-02	3,29E-02	5,05	1,00	0,85	4,32	1,51E-01	1,12E-01	5,05
20,79	1,00	0,93	4,69	6,39E-02	5,16E-02	5,06	1,00	0,82	4,17	2,34E-01	1,74E-01	5,06
25,40	1,00	0,91	4,63	9,57E-02	7,90E-02	5,07	1,00	0,79	4,03	3,41E-01	2,50E-01	5,08
30,31	1,00	0,90	4,56	1,37E-01	1,07E-01	5,08	1,00	0,76	3,89	4,73E-01	3,40E-01	5,09
35,50	1,00	0,88	4,49	1,88E-01	1,45E-01	5,10	1,00	0,73	3,75	6,32E-01	4,53E-01	5,11
40,95	1,00	0,86	4,42	2,50E-01	2,05E-01	5,11	1,00	0,71	3,62	8,18E-01	5,94E-01	5,13
46,66	1,00	0,85	4,35	3,26E-01	2,45E-01	5,13	1,00	0,68	3,49	1,03E+00	7,46E-01	5,15
52,62	1,00	0,83	4,28	4,15E-01	3,28E-01	5,15	1,00	0,65	3,37	1,27E+00	9,03E-01	5,17
58,80	1,00	0,81	4,21	5,19E-01	4,10E-01	5,17	1,00	0,63	3,26	1,54E+00	1,08E+00	5,19
65,21	1,00	0,80	4,14	6,40E-01	5,00E-01	5,19	1,00	0,60	3,15	1,84E+00	1,28E+00	5,21
71,84	1,00	0,78	4,07	7,77E-01	6,24E-01	5,21	1,00	0,58	3,05	2,17E+00	1,50E+00	5,23
78,68	1,00	0,76	3,99	9,33E-01	7,11E-01	5,23	1,00	0,56	2,95	2,52E+00	1,68E+00	5,26
85,73	1,00	0,75	3,92	1,11E+00	8,54E-01	5,25	1,00	0,54	2,86	2,90E+00	1,99E+00	5,28
92,97	1,00	0,73	3,85	1,31E+00	9,93E-01	5,27	1,00	0,52	2,77	3,31E+00	2,21E+00	5,30
100,40	1,00	0,72	3,78	1,52E+00	1,11E+00	5,29	1,00	0,50	2,69	3,74E+00	2,46E+00	5,33

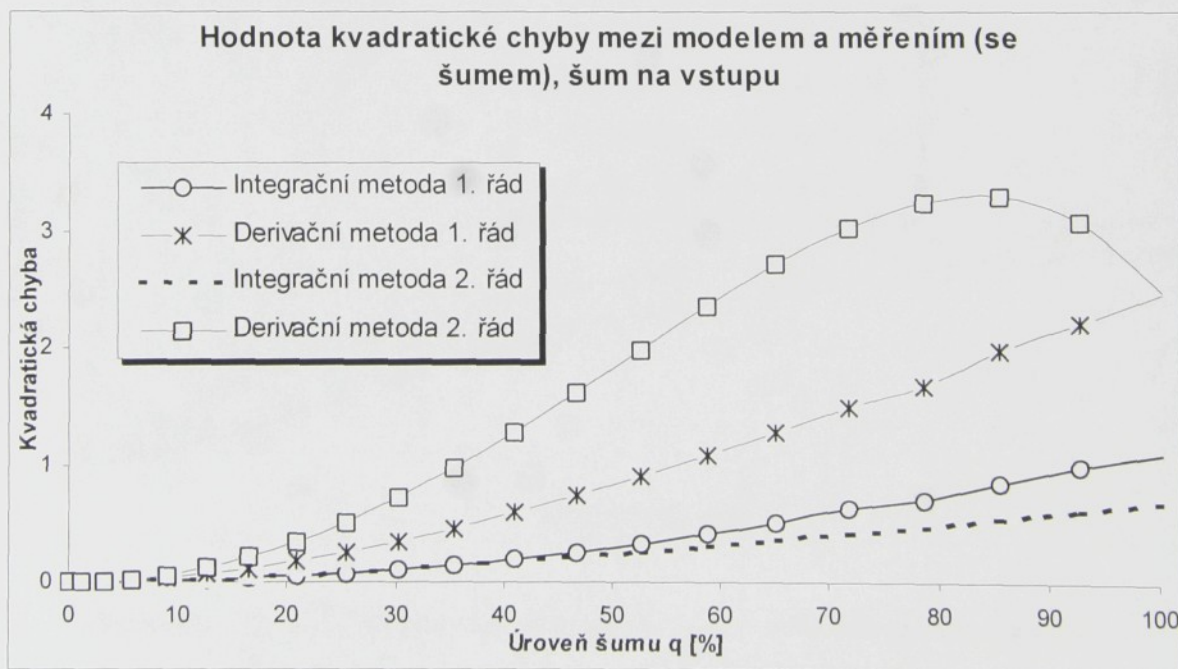
tabulka III.9-3: Vliv poruch na vstupu na derivační a integrační metodu pro soustavu 1. řádu



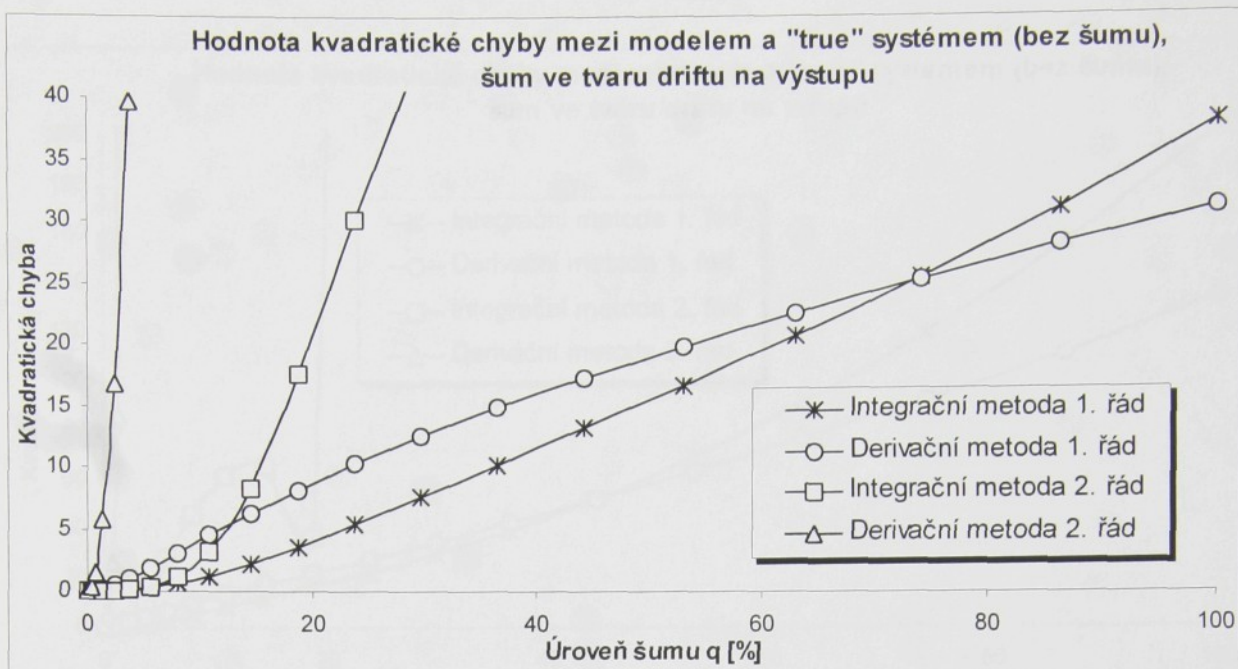
obrázek III.9-4: Kvadratická chyba J₁ v závislosti na úrovni poruchy působící na vstupu

q	a ₂	a ₁	a ₀	b	J ₁	J ₂	k	ad ₂	ad ₁	ad ₀	bd	J ₁	J ₂	kd
0,00	1,00	1,00	1,00	5,00	-5,56E-05	-5,56E-05	5,00	1,00	1,00	1,00	4,99	-2,49E-06	-2,49E-06	5,00
1,12	1,00	1,01	1,00	5,03	3,83E-04	1,59E-04	5,00	1,00	1,01	1,01	5,04	5,51E-04	6,09E-04	5,00
3,18	1,00	1,03	1,01	5,07	3,32E-03	1,53E-03	5,01	1,00	1,02	1,03	5,14	6,35E-03	6,73E-03	5,01
5,83	1,00	1,05	1,02	5,13	1,11E-02	5,14E-03	5,02	1,00	1,04	1,05	5,29	2,34E-02	2,45E-02	5,02
8,98	1,00	1,08	1,04	5,20	2,59E-02	1,18E-02	5,03	1,00	1,07	1,09	5,47	5,82E-02	6,01E-02	5,03
12,55	1,00	1,11	1,05	5,29	4,96E-02	2,22E-02	5,04	1,00	1,10	1,13	5,69	1,18E-01	1,20E-01	5,04
16,50	1,00	1,14	1,06	5,37	8,37E-02	3,67E-02	5,05	1,00	1,15	1,18	5,97	2,09E-01	2,09E-01	5,05
20,79	1,00	1,18	1,08	5,47	1,30E-01	5,57E-02	5,06	1,00	1,20	1,24	6,30	3,41E-01	3,35E-01	5,06
25,40	1,00	1,22	1,10	5,57	1,88E-01	7,94E-02	5,08	1,00	1,27	1,32	6,72	5,20E-01	5,03E-01	5,08
30,31	1,00	1,26	1,12	5,68	2,60E-01	1,07E-01	5,09	1,00	1,35	1,42	7,23	7,55E-01	7,15E-01	5,09
35,50	1,00	1,31	1,13	5,79	3,47E-01	1,40E-01	5,11	1,00	1,46	1,54	7,87	1,05E+00	9,73E-01	5,11
40,95	1,00	1,35	1,15	5,90	4,47E-01	1,76E-01	5,12	1,00	1,59	1,70	8,70	1,41E+00	1,28E+00	5,13
46,66	1,00	1,40	1,17	6,01	5,62E-01	2,16E-01	5,14	1,00	1,78	1,90	9,78	1,83E+00	1,62E+00	5,15
52,62	1,00	1,45	1,18	6,11	6,90E-01	2,60E-01	5,16	1,00	2,03	2,17	11,23	2,31E+00	1,98E+00	5,17
58,80	1,00	1,49	1,20	6,20	8,33E-01	3,07E-01	5,17	1,00	2,39	2,55	13,25	2,82E+00	2,36E+00	5,19
65,21	1,00	1,54	1,21	6,28	9,88E-01	3,58E-01	5,19	1,00	2,93	3,11	16,19	3,34E+00	2,72E+00	5,21
71,84	1,00	1,58	1,22	6,34	1,16E+00	4,14E-01	5,21	1,00	3,78	3,96	20,71	3,82E+00	3,03E+00	5,23
78,68	1,00	1,61	1,22	6,37	1,34E+00	4,72E-01	5,23	1,00	5,18	5,32	27,93	4,22E+00	3,24E+00	5,25
85,73	1,00	1,63	1,21	6,36	1,53E+00	5,38E-01	5,25	1,00	7,46	7,40	39,02	4,46E+00	3,29E+00	5,27
92,97	1,00	1,65	1,20	6,32	1,73E+00	6,09E-01	5,27	1,00	9,83	9,21	48,77	4,46E+00	3,08E+00	5,30
100,40	1,00	1,65	1,18	6,23	1,94E+00	6,87E-01	5,29	1,00	8,48	7,23	38,51	4,12E+00	2,49E+00	5,32

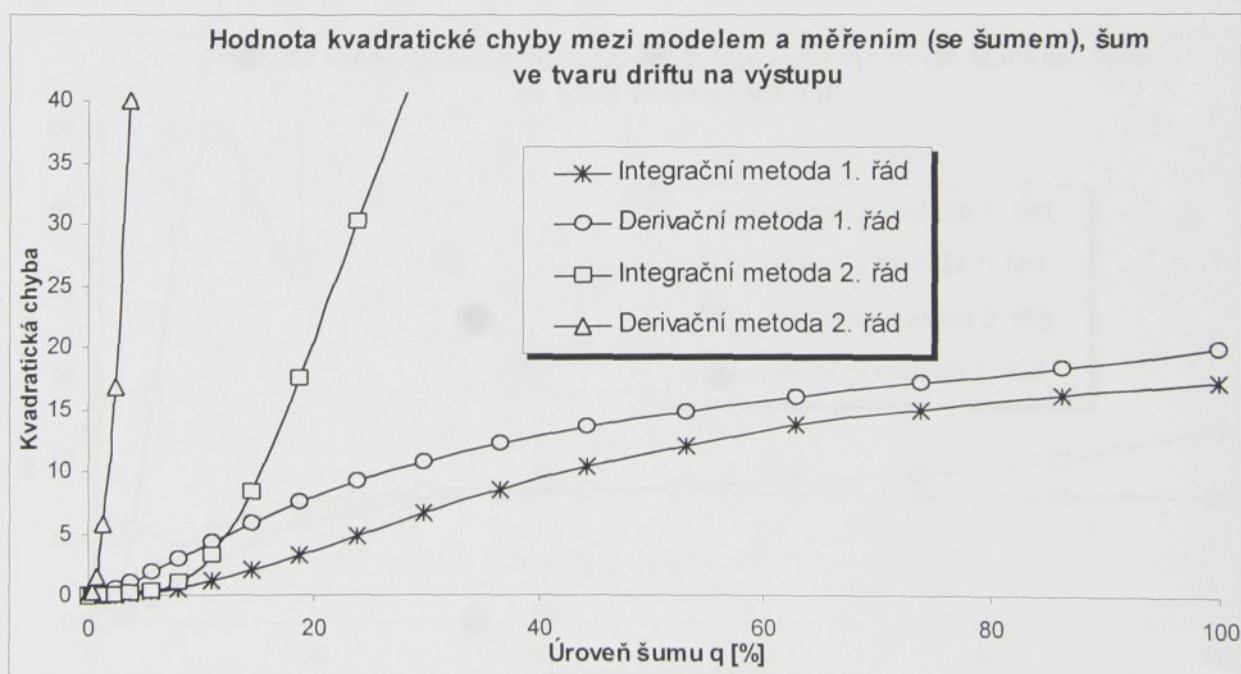
tabulka III.9-4: Vliv poruch na vstupu na derivační a integrační metodu pro soustavu 2. řádu



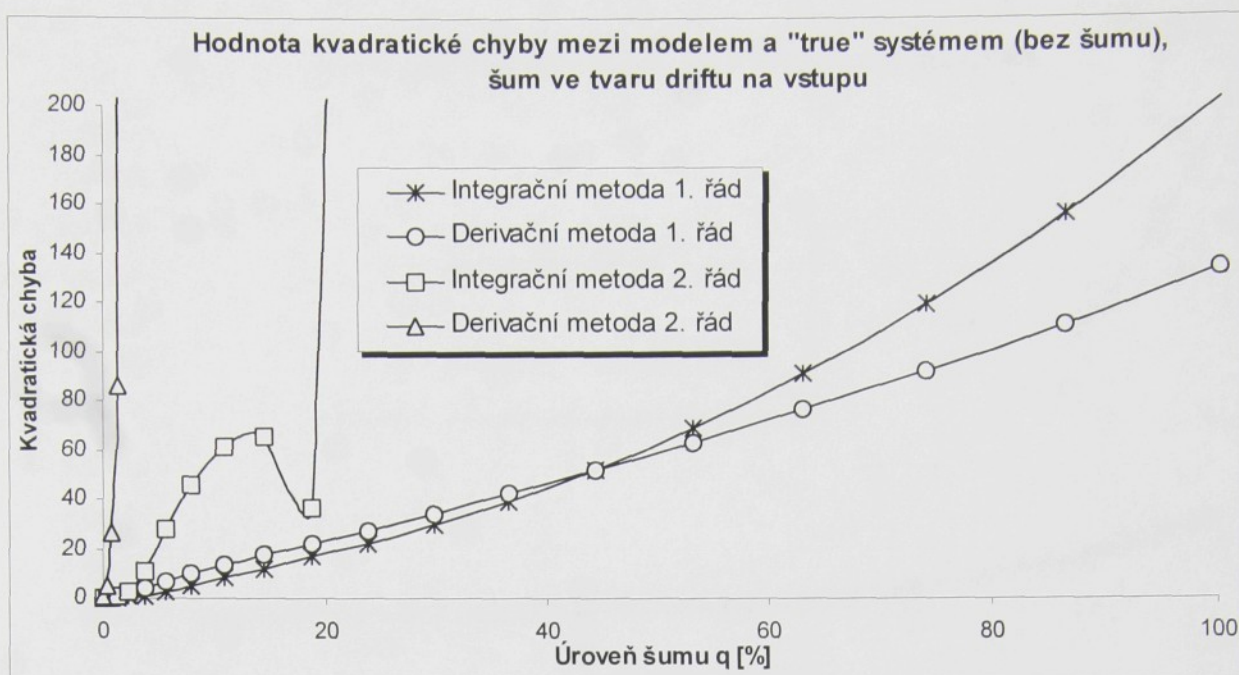
obrázek III.9-5: Kvadratická chyba J₂ v závislosti na úrovni poruchy působící na vstupu



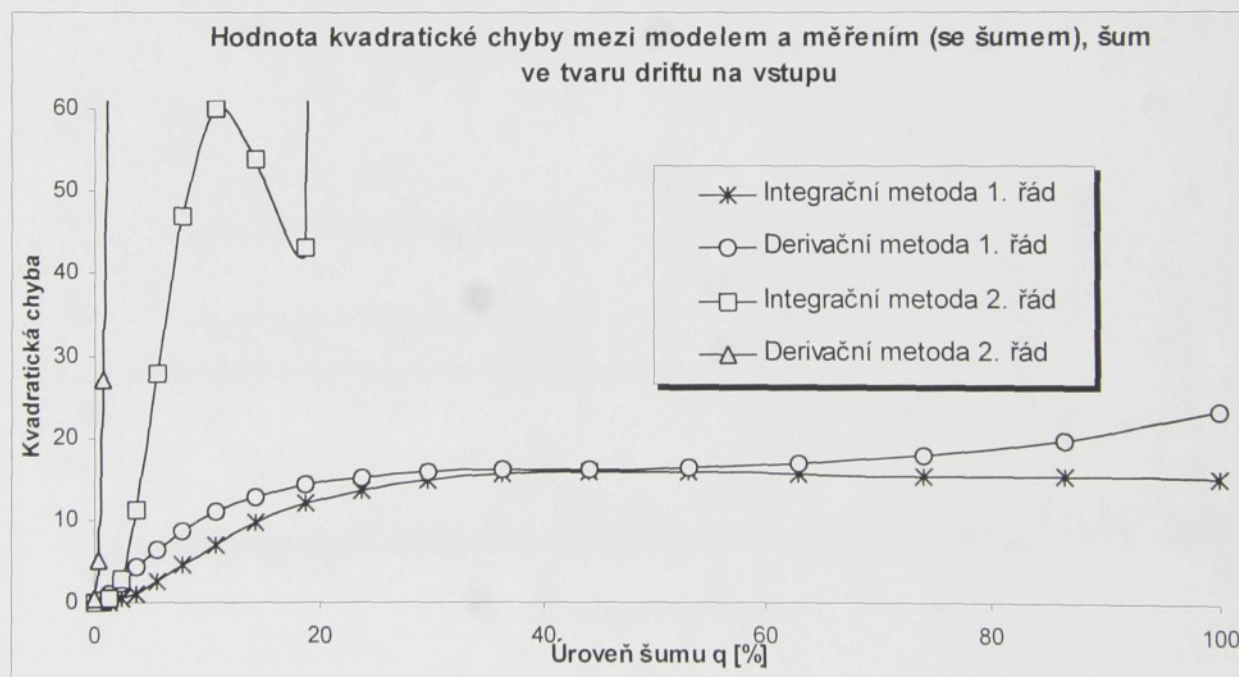
obrázek III.9-6: Kvadratická chyba J_1 v závislosti na úrovni poruchy ve tvaru driftu působící na výstupu



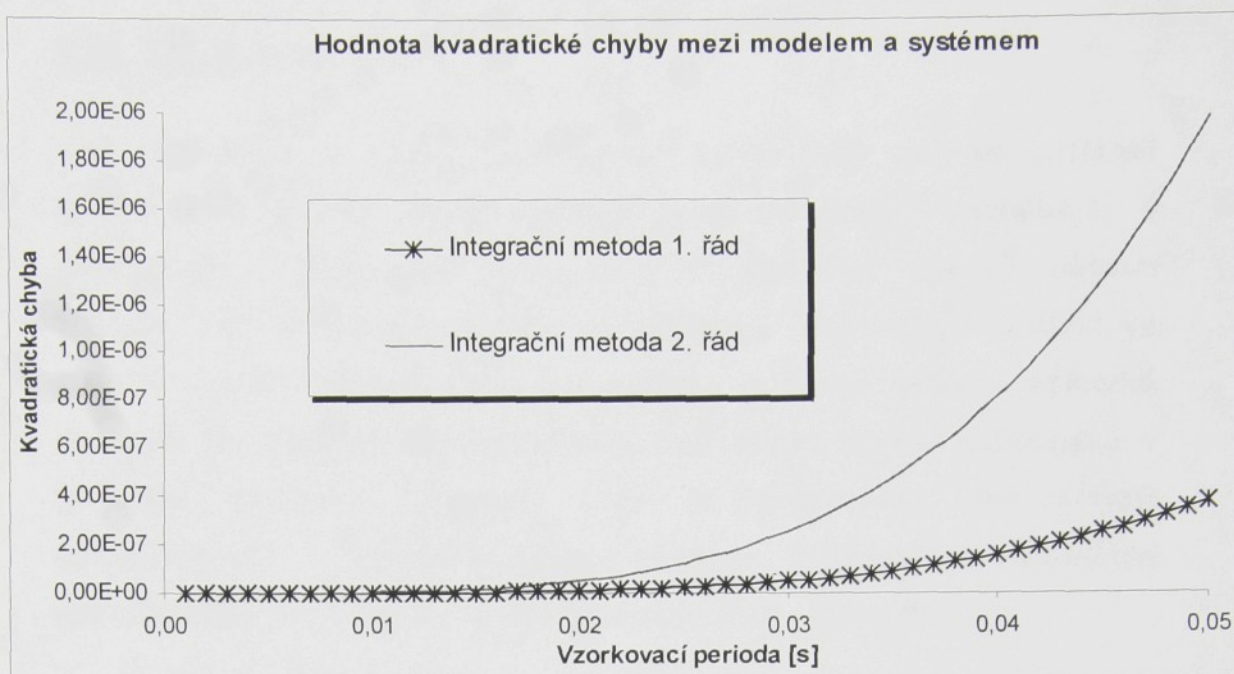
obrázek III.9-7: Kvadratická chyba J_2 v závislosti na úrovni poruchy ve tvaru driftu působící na výstupu



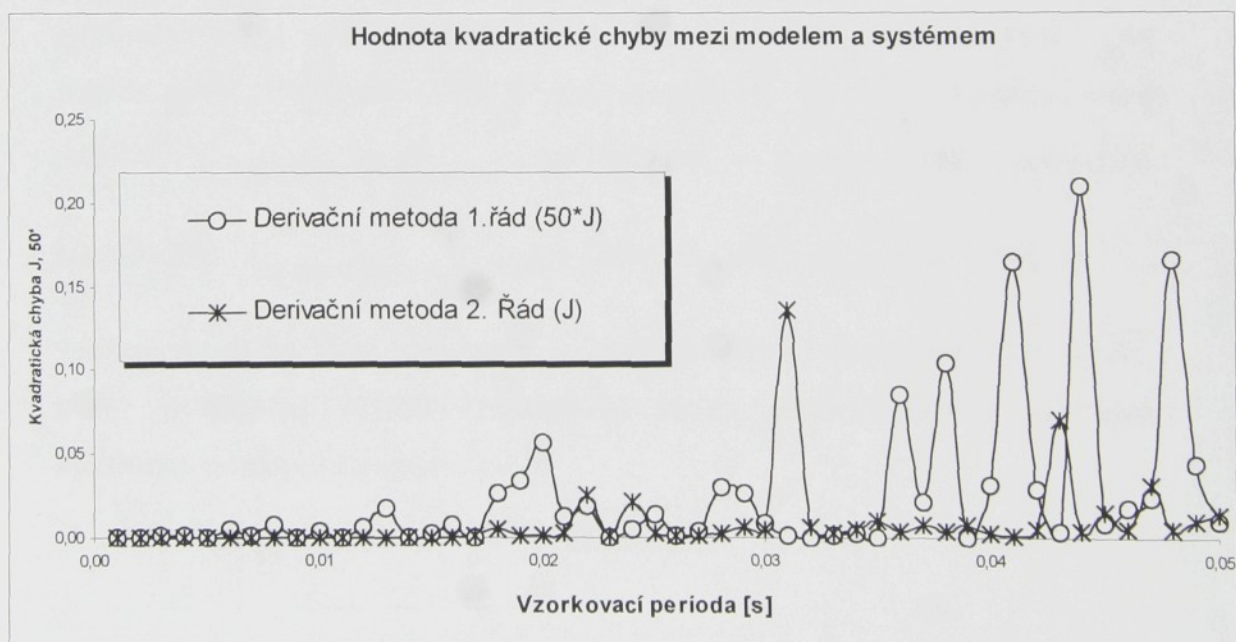
**obrázek III.9-8: Kvadratická chyba J_1 v závislosti na úrovni
poruchy ve tvaru driftu působící na vstupu**



**obrázek III.9-9: Kvadratická chyba J_2 v závislosti na úrovni
poruchy ve tvaru driftu působící na vstupu**



obrázek III.9-10: Kvadratická chyba J_1 v závislosti na velikosti vzorkovací periody pro integrační metodu



obrázek III.9-11: Kvadratická chyba J_1 v závislosti na velikosti vzorkovací periody pro derivační metodu

III.10 ZÁVĚR

Závěrem je nutno podotknout, že metody pro identifikaci modelů ve spojité časové oblasti jsou numericky náročnější v porovnání s metodami pracujícími v diskrétní časové oblasti [LJU2]. Tato práce ale ukazuje, že přístup k identifikaci modelů ve spojité časové oblasti (tedy té reálné, ve které většina systémů pracuje) je výhodný především tam, kde máme nějaké informace o podstatě takového procesu (děj je popsitelný fyzikálními zákonitostmi) a hledáme pouze skupinu neznámých parametrů specifických pro daný konkrétní problém (tzv. "grey box").

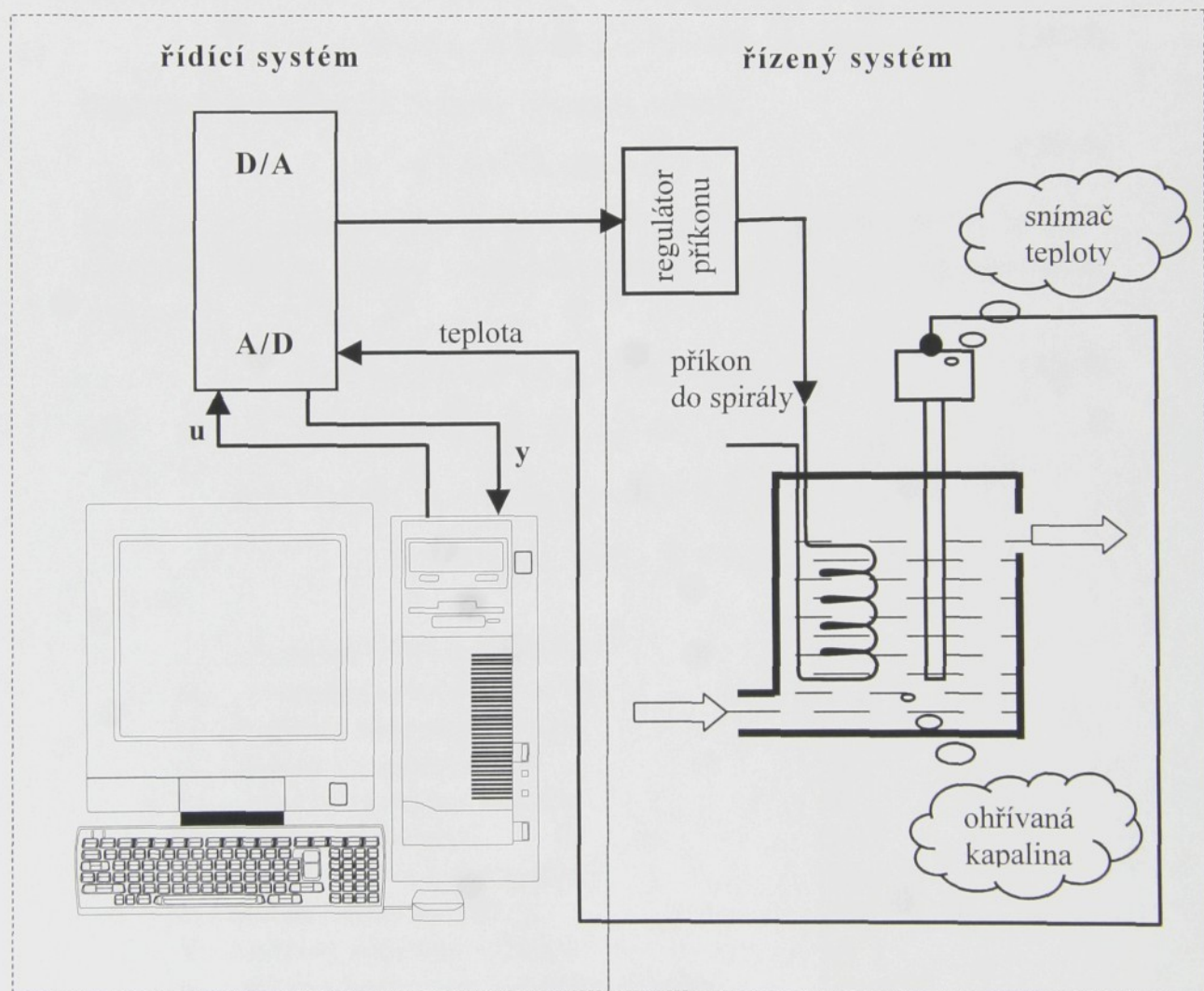
Metody identifikace modelů lineárních v parametrech tak, jak byly prezentovány v kapitolách III.2 až III.5, jsou vhodné především pro inicializační odhad parametrů, který je možné použít pro některou obecnou optimalizační metodu. A to z důvodu, že při řešení nahrazujeme vnitřní proměnné modelu (který je nelineární v parametrech) korespondujícími hodnotami změřenými na studovaném systému. Např. pro soustavu 1. řádu nahrazujeme model: $y_m(t, \mathbf{p}) = b \int_0^t u(t) - a \int_0^t y_m(t, \mathbf{p})$ (nelineární v parametrech) modelem lineárním: $y_m(t, \mathbf{p}) = b \int_0^t u(t) - a \int_0^t y(t)$, $\mathbf{p} = (a, b)$, y značí změřené hodnoty, y_m výstup modelu. Tato okolnost, v případě silně zašuměného systému, může podstatně zhoršit schopnost těchto metod nalézt globální optimum účelové funkce.

IV Řešené úlohy

Tato kapitola, která uzavírá předkládanou práci, dokumentuje použitelnost definovaných metod hledání modelů dynamických systémů z experimentální dat ve spojitě časové oblasti na řadě skutečných aplikačních příkladů. Různé experimenty byly provedeny na čtyřech studovaných systémech a změřená vstupně-výstupní data jsou analyzována metodami identifikace ve spojitě časové oblasti. Struktura modelu je nalezena pomocí matematicko-fyzikální analýzy, přičemž je v některých případech zjednodušena. Nalezené modely jsou následovně podrobeny kvalitativní analýze a jednotlivé metody jsou porovnány mezi sebou. První studovanou soustavu tvoří průtokový ohřívač (kapitola IV.1). Následující dva systémy jsou stejnosměrné motorky. První z nich je stejnosměrný motor s permanentními magnety (kapitola IV.2), druhý je stejnosměrný motor s pružnou spojkou (kapitola IV.3). Poslední studovaný systém představuje zásobník kapaliny, ve které se reguluje výška hladiny (kapitola IV.4).

IV.1 IDENTIFIKACE PRŮTOKOVÉHO VÝMĚNÍKU

Zaměřme nyní naši pozornost na systém s principiálním schématem zobrazeném na následujícím obrázku (obrázek IV.1-1). Jedná se o počítačem řízený průtokový výměník, do něhož vstupuje konstantní průtokové množství pracovní kapaliny, která je ohřívána elektrickou spirálou (topným článkem) s regulovatelným příkonem. Styk počítače a studovaného systému je realizován pomocí převodníkové karty. Teplota média je snímána odporovým snímačem



obrázek IV.1-1: Průtokový výměník - schéma systému

teploty umístěném uvnitř výměníku a tento signál je připojen na analogový vstup převodníkové karty. Výstupy z převodníkové karty

jsou připojeny k tyristorovému regulátoru příkonu do topného článku.

Nyní se pokusme nalézt diferenciální rovnici popisující dynamiku teploty ve výměníku v závislosti na příkonu topného článku a teplotě vstupujícího média. Předpokládejme, že teplota uvnitř výměníku v [$^{\circ}\text{C}$] je homogenní.

Tepelná kapacita výměníku je dána vztahem:

$$W = V\rho c = \kappa_0 v, \quad k_0 = V\rho c. \quad (\text{IV-1})$$

Obdobně pro topný článek:

$$W_1 = V_1\rho_1 c_1 = \kappa_1 v_1, \quad k_1 = V_1\rho_1 c_1. \quad (\text{IV-2})$$

Energetická bilance ve výměníku vede na rovnost:

$$k_2(v_1 - v) = k_0\dot{v} + k_3(v - v_0), \quad k_2 = \alpha S_1, \quad k_3 = Q\rho c. \quad (\text{IV-3})$$

Obdobně energetická bilance topného tělesa:

$$q = k_1\dot{v} + k_2(v_1 - v). \quad (\text{IV-4})$$

Po úpravě z rovnic (IV-1) až (IV-4) získáme diferenciální rovnici druhého řádu s dvěma vstupy a jedním výstupem ve standardním tvaru:

$$a_2\ddot{y} + a_1\dot{y} + a_0y = b_0u_1 + c_0u_2 + c_1\dot{u}_2, \quad (\text{IV-5})$$

kde:

$$y = k_{AD}v, \quad u_1 = k_{DA}q, \quad u_2 = v_0,$$

$$a_0 = k_{AD}k_2k_3, \quad a_1 = k_{AD}[k_1(k_2 + k_3) + k_0k_2], \quad a_2 = k_{AD}k_0k_1,$$

$$c_0 = k_2k_3, \quad c_1 = k_1k_3, \quad b_0 = k_{DA}k_2,$$

v ...výstupní teplota kapaliny	[$^{\circ}\text{C}$]
v_0 ...vstupní teplota kapaliny	[$^{\circ}\text{C}$]
v_1 ...teplota topného článku	[$^{\circ}\text{C}$]
q ...příkon výměníku	[W]
S_1 ...plocha topného článku	[m^2]
V ...objem výměníku	[m^3]
ρ ...měrná hmotnost kapaliny	[kgm^{-3}]
c ...měrné teplo kapaliny	[$\text{Jkg}^{-1}\text{K}^{-1}$]
V_1 ...objem topného článku	[m^3]
ρ_1 ...měrná hmotnost topného článku	[kgm^{-3}]
c_1 ...měrné teplo topného článku	[$\text{Jkg}^{-1}\text{K}^{-1}$]
α ...koeficient přestupu tepla	[$\text{WK}^{-1}\text{m}^{-2}$]
Q ...průtok výměníkem	[m^3s^{-1}]
k_{AD} ...zesílení AD převodníku	[1]
k_{DA} ...zesílení DA převodníku	[1]

V průběhu experimentu se teplota vstupující kapaliny neměnila. Proto můžeme získaný model (IV-5) ještě zjednodušit na systém s jedním vstupem. Změřená data byla dále upravena posunem počátku souřadného systému do počátečního ustáleného bodu (místo absolutní teploty uvažujeme pouze rozdíl teplot). Model byl tedy uvažován ve tvaru:

$$\ddot{y} + a_1\dot{y} + a_0y = b_0u_1. \tag{IV-6}$$

Nalezené parametry integrační metodou:

$$a_1 = 0.019656, a_0 = 0.00020399, \\ b_0 = 3.953084e-005.$$

Nalezené parametry derivační metodou:

$$a_1 = 0.019122, a_0 = 0.00019798, \\ b_0 = 3.836786e-005.$$

Parametry nalezené těmito metodami jsou velice podobné a i, jak bylo vyzkoušeno, odezvy těchto modelů jsou téměř identické. Stejně je tomu pak i s následným vyhodnocením kvality modelu (nezaznamenali jsme podstatné rozdíly), proto jsou zde uvedeny pouze výsledky srovnání modelu a systému pro model nalezený integrační metodou (obrázek IV.1-2 - obrázek IV.1-6).

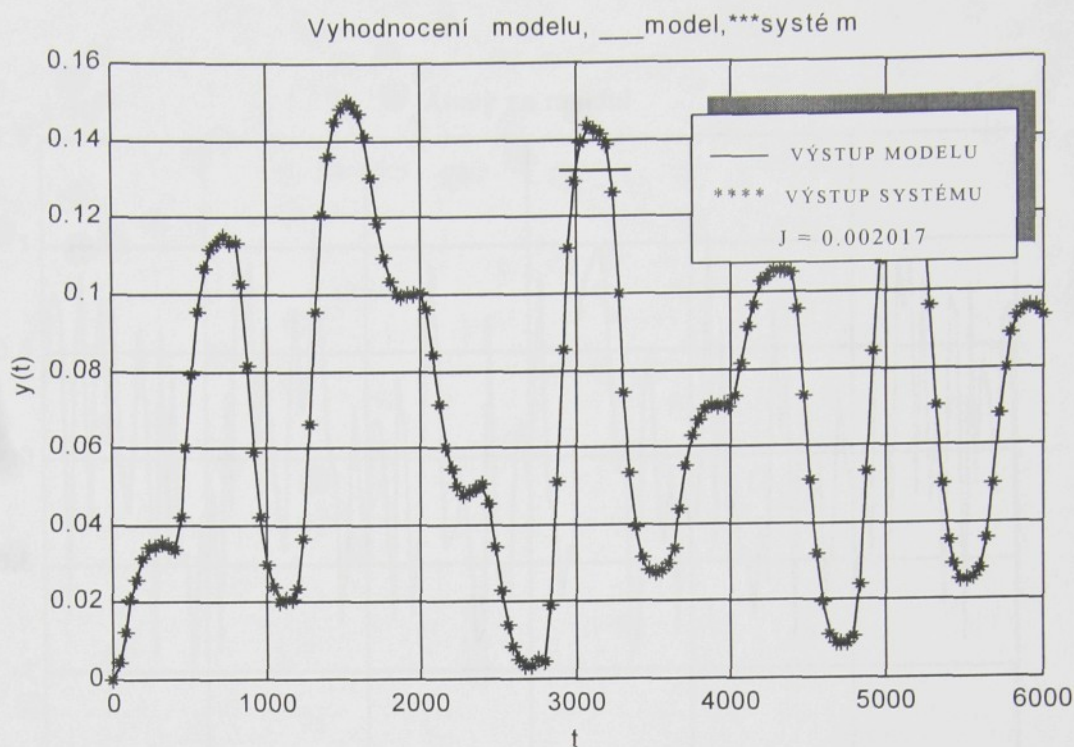
Nakonec ještě vyzkoušejme nalézt parametry stejné struktury iterační metodou (použita simplexová optimalizační metoda), kterou inicializujeme parametry získanými pomocí derivační metody.

Nalezené parametry iterační metodou:

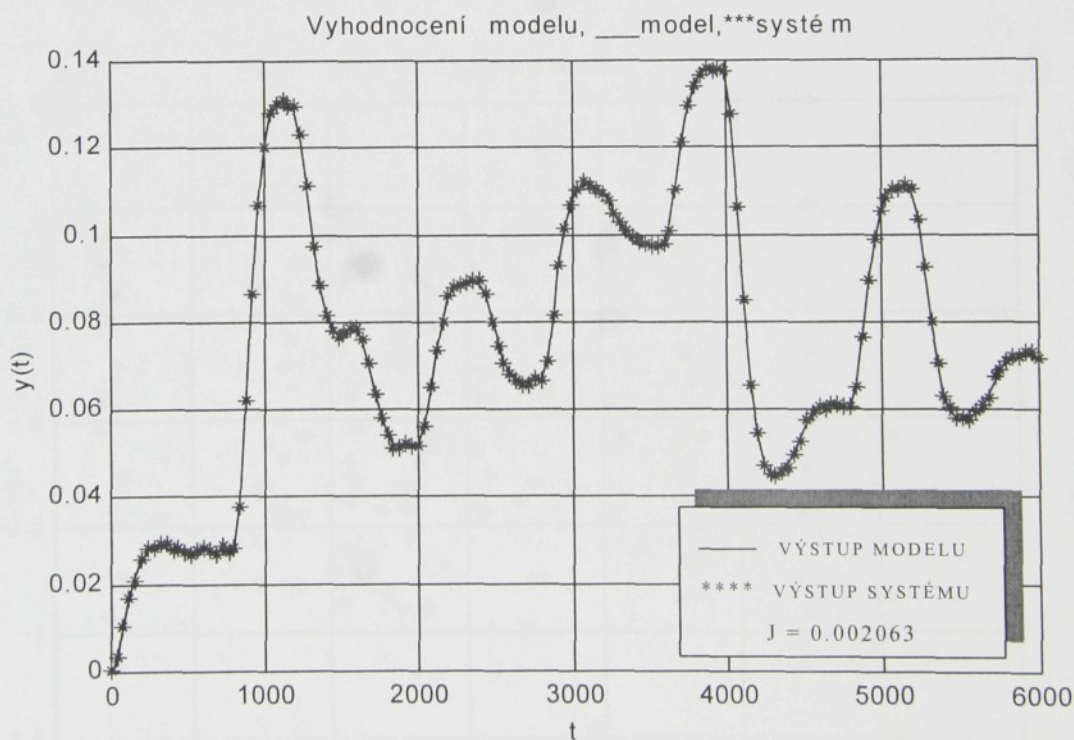
$$a_2 = 1, a_1 = 0.019426, a_0 = 0.0002011 \\ b_0 = 3.894428e-005.$$

<i>Metoda</i>	<i>Hodnota účelové funkce</i> <i>J(p)</i>
<i>Integrační metoda</i>	0.002017
<i>Derivační metoda</i>	0.001923
<i>Iterační metoda</i>	0.001776

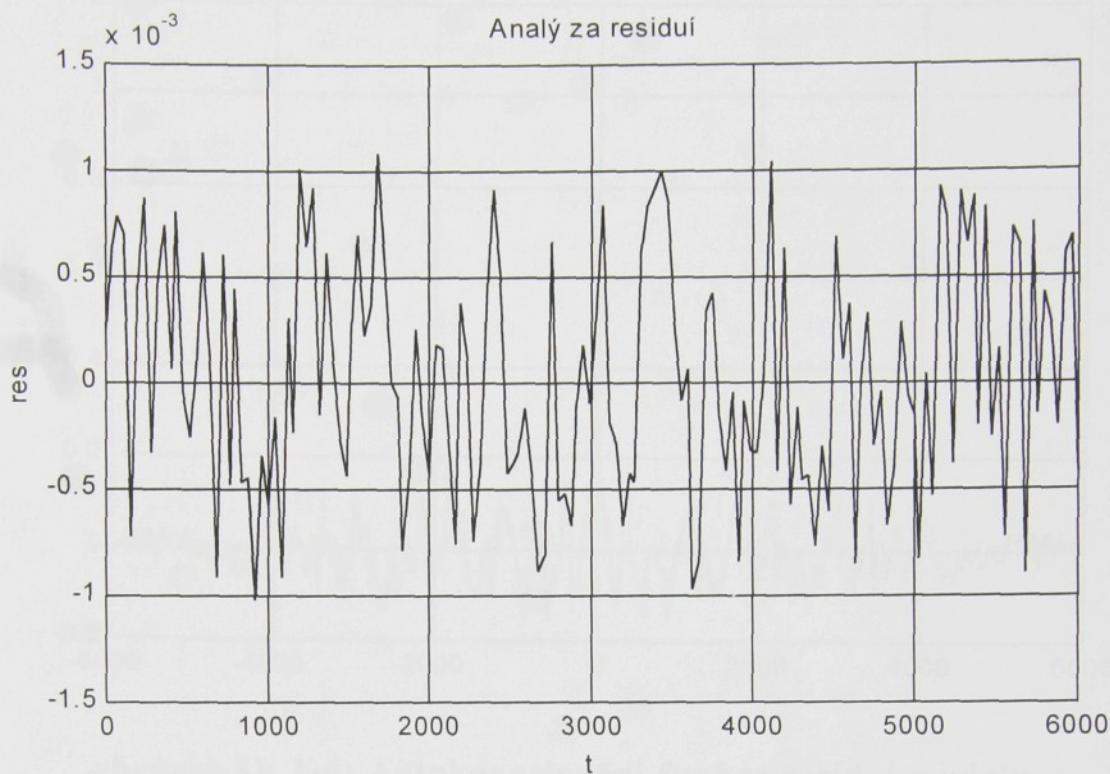
tabulka IV.1-1: Hodnoty účelové funkce pro systém výměníku



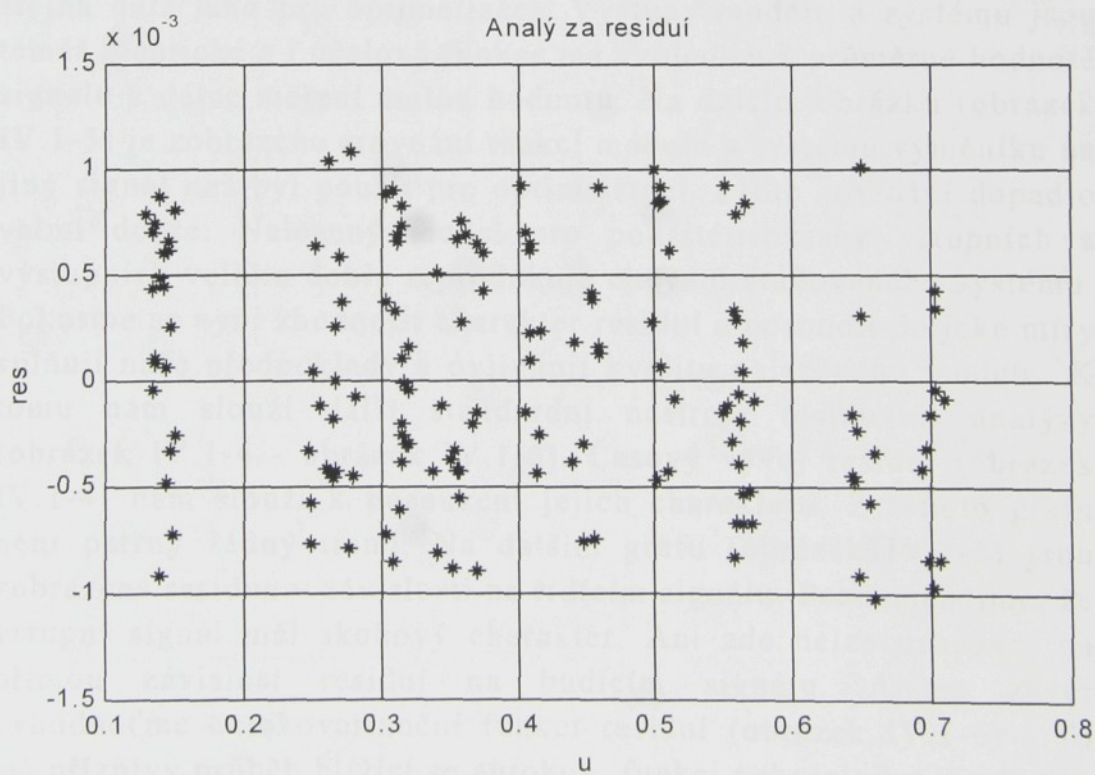
**obrázek IV.1-2: Porovnání výstupu modelu a systému výměníku
na datech použitých pro optimalizaci**



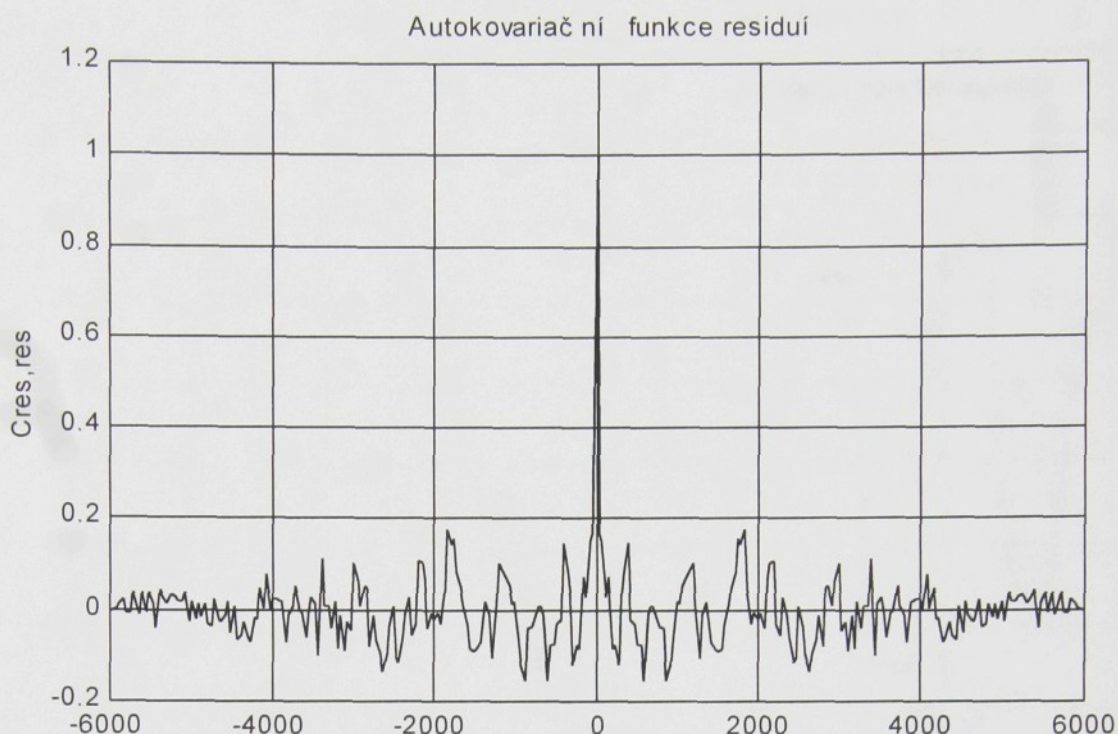
**obrázek IV.1-3: Porovnání výstupu modelu a systému výměníku
na jiných datech než byla použita pro optimalizaci**



obrázek IV.1-4: Časová posloupnost residuí modelu výměníku

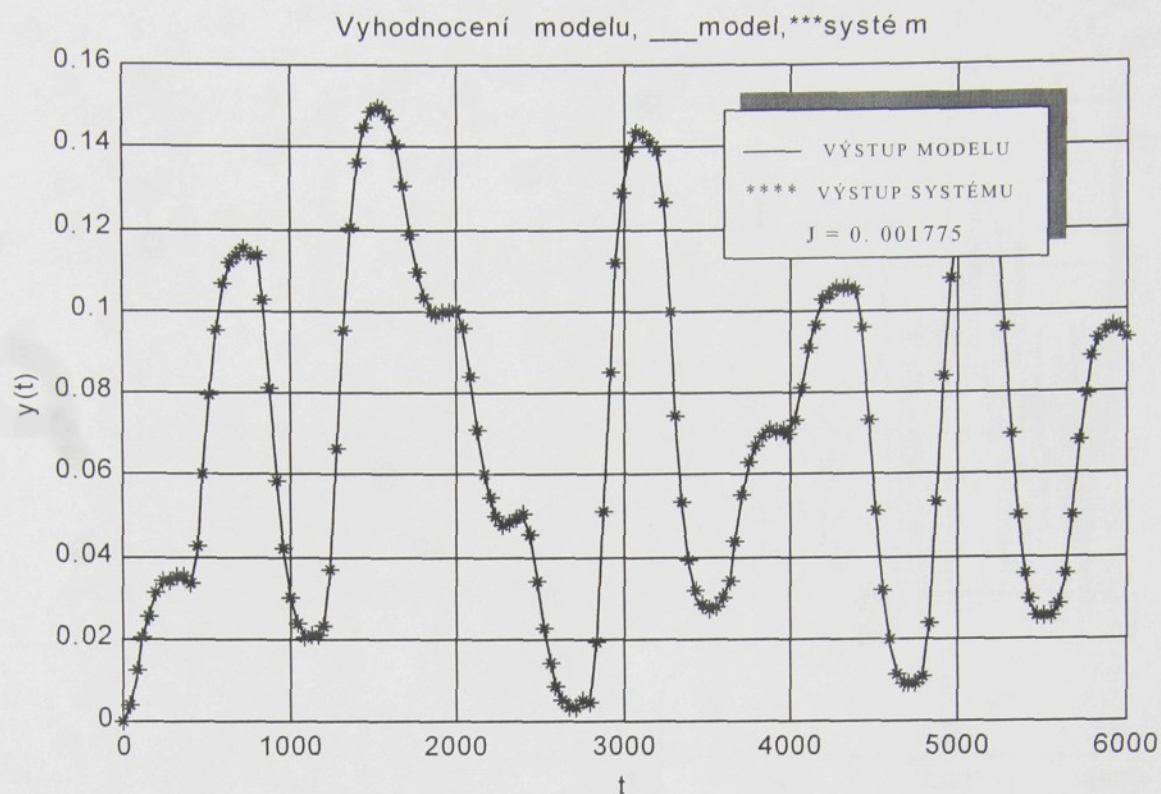


obrázek IV.1-5: Graf residuí v závislosti na budícím signálu

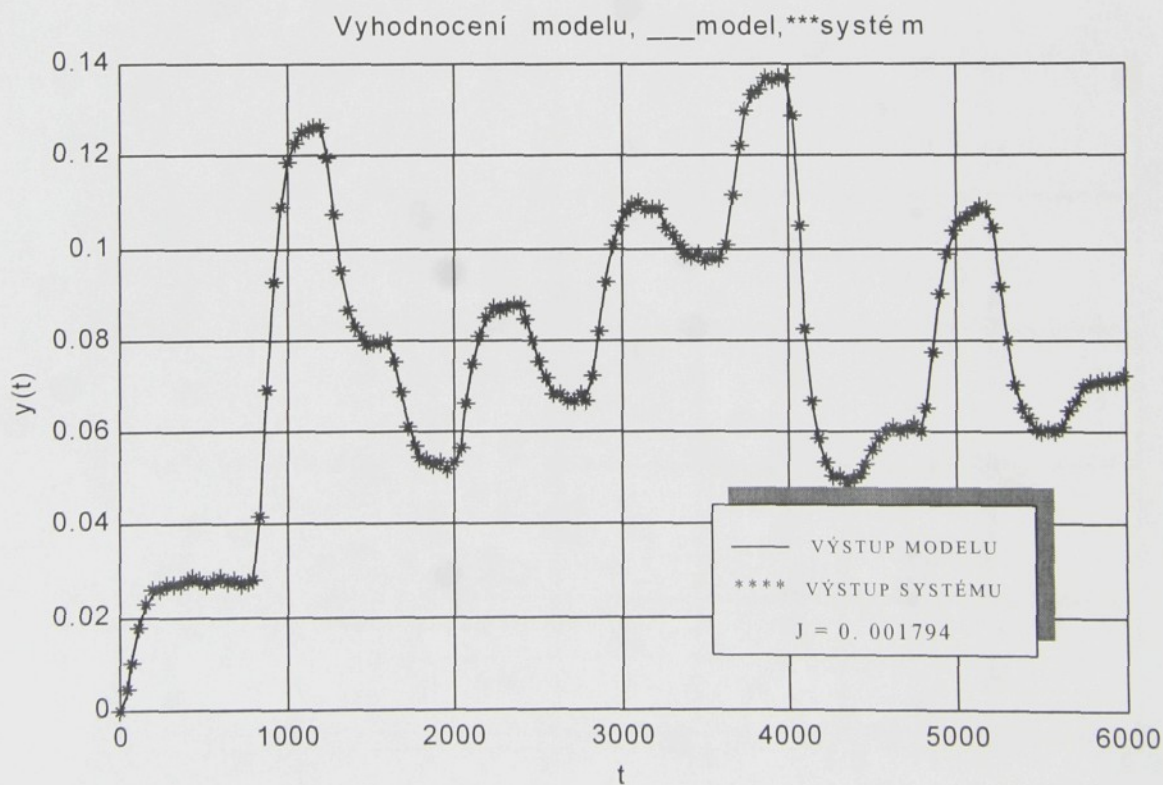


**obrázek IV.1-6: Autokovarianční funkce residuí modelu
výměníku**

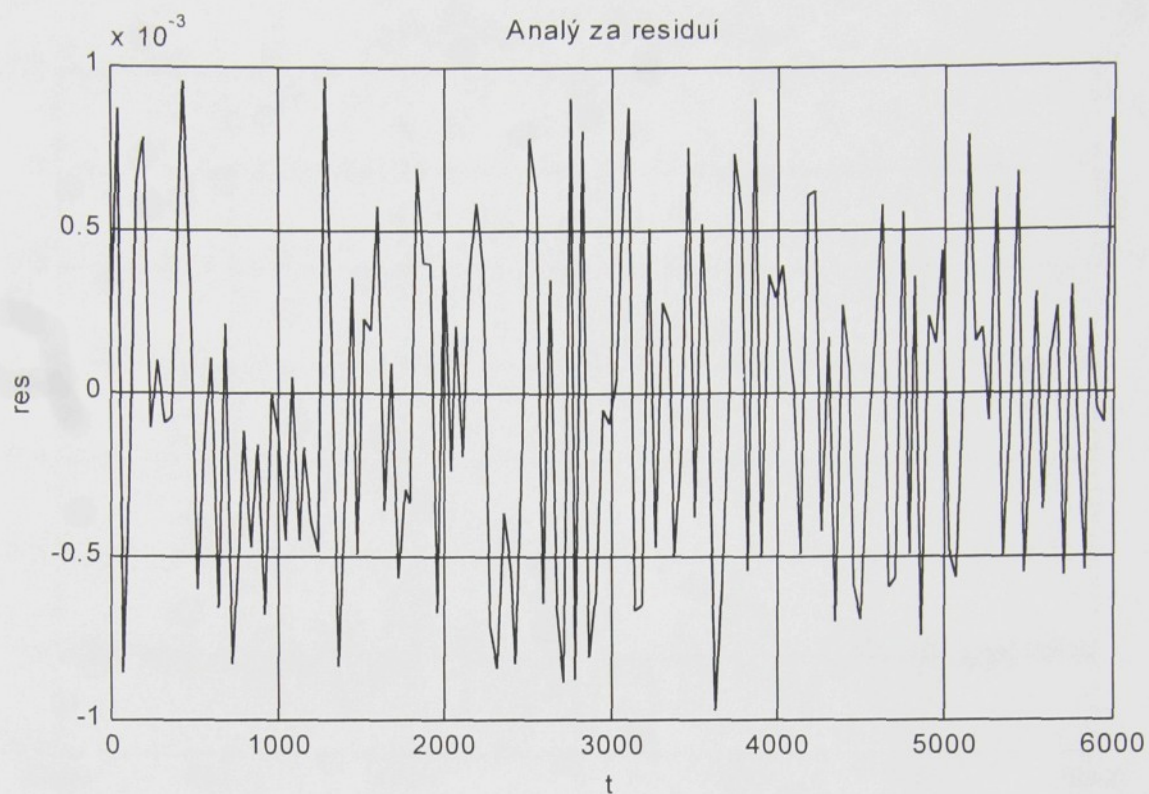
Podívejme se na porovnání výstupu modelu a vyšetřovaného systému výměníku (obrázek IV.1-2). Pro toto srovnání byla použita stejná data jako pro optimalizaci. Výstupy modelu a systému jsou téměř identické a i účelová funkce má vzhledem k průměrné hodnotě signálu a délce měření malou hodnotu. Na dalším obrázku (obrázek IV.1-3) je zobrazeno srovnání reakcí modelu a systému výměníku na jiný signál než byl použit pro optimalizaci. I toto srovnání dopadlo velmi dobře. Nalezený model pro použité rozsahy vstupních a výstupních veličin dobře reprodukuje chování studovaného systému. Pokusme se nyní zhodnotit charakter residuí a posoudit do jaké míry splňují naše předpoklady a ovlivňují kvalitu nalezeného modelu. K tomu nám slouží další standardní nástroje residuální analýzy (obrázek IV.1-4 - obrázek IV.1-6). Časový vývoj residuí (obrázek IV.1-4) nám slouží k posouzení jejich charakteru. Z tohoto grafu není patrný žádný trend. Na dalším grafu (obrázek IV.1-5) jsou zobrazena residua v závislosti na řídicím signálu. Poznamenejme, že vstupní signál měl skokový charakter. Ani zde nelze usuzovat na přímou závislost residuí na budícím signálu. A na závěr vyhodnotíme autokovarianční funkci residuí (obrázek IV.1-6). Ta má příznivý průběh blížíci se autokov. funkci nekorelovaného šumu. Na závěr lze říci, že tento model prošel všemi uvažovanými testy a je tedy vhodný pro použití v oblasti automatického řízení.



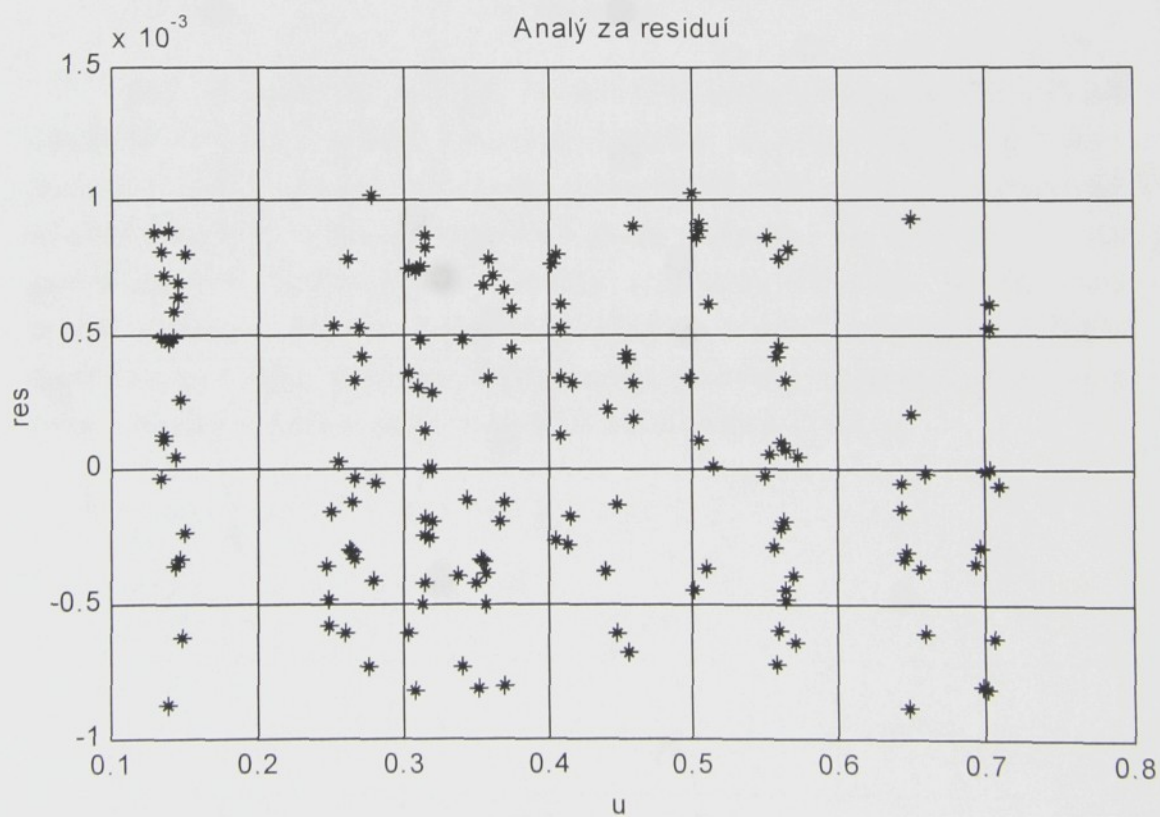
**obrázek IV.1-7: Porovnání výstupu modelu a systému výměníku
na datech použitých pro optimalizaci**



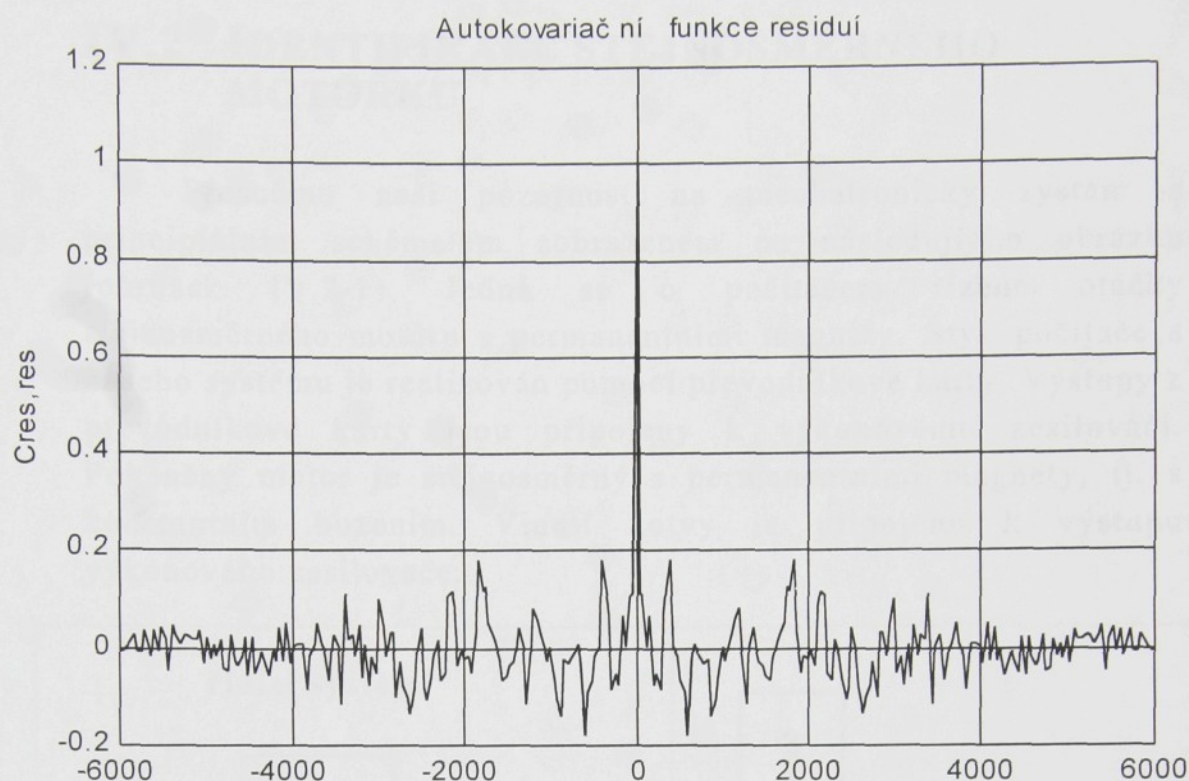
**obrázek IV.1-8: Porovnání výstupu modelu a systému výměníku
na jiných datech než byla použita pro optimalizaci**



obrázek IV.1-9: Časová posloupnost residuí modelu výměníku



obrázek IV.1-10: Graf residuí v závislosti na budícím signálu

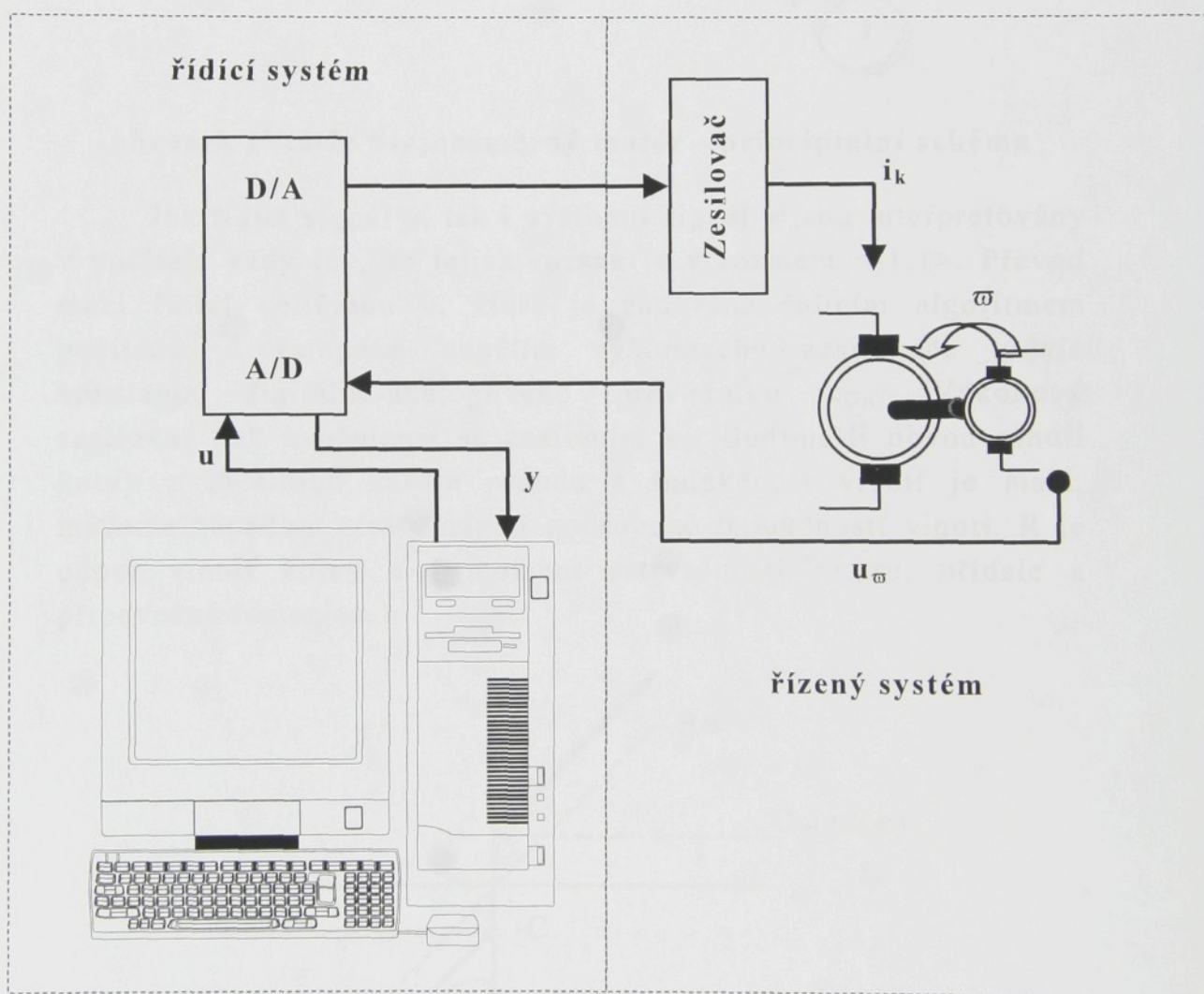


**obrázek IV.1-11: Autokovarianční funkce residuí modelu
výměníku**

Jak je vidět z analýzy kvality modelu (obrázek IV.1-7 až obrázek IV.1-11), model nalezený iterační metodou se chová velmi podobně jako model nalezený integrační metodou. Kvadratická účelová funkce vykazuje poněkud nižší hodnoty. To se projeví i na grafu časové posloupnosti residuí (obrázek IV.1-9), které mají menší rozptyl. Jak se dalo předpokládat, model nalezený iterační metodou poněkud přesněji reprodukuje chování systému, než tomu bylo u modelu nalezeného metodou nejmenších čtverců.

IV.2 IDENTIFIKACE STEJNOSMĚRNÉHO MOTORKU

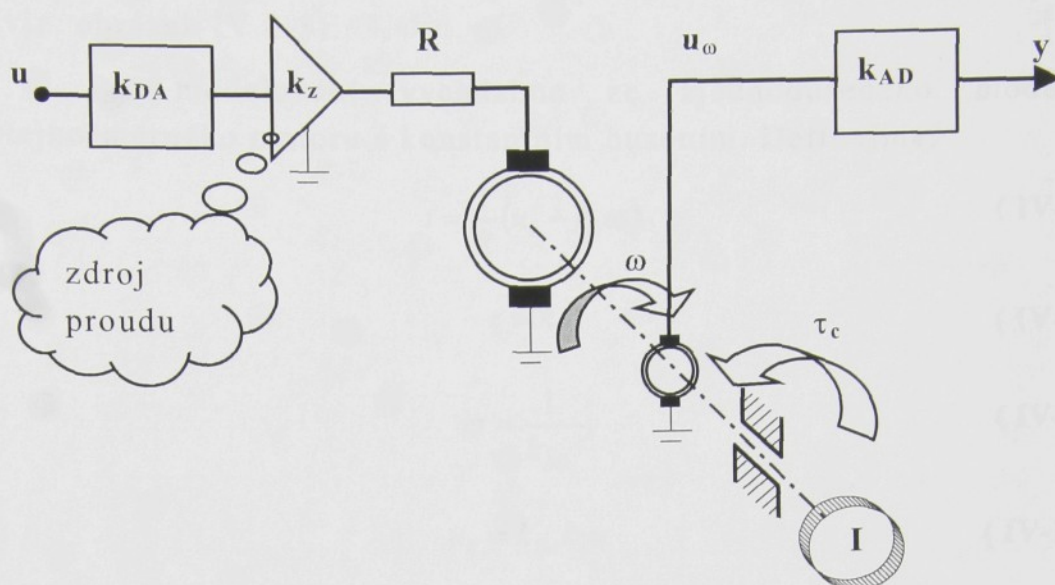
Přesuňme naši pozornost na mechatronický systém s principiálním schématem zobrazeném na následujícím obrázku (obrázek IV.2-1). Jedná se o počítačem řízené otáčky stejnosměrného motoru s permanentními magnety. Styk počítače a našeho systému je realizován pomocí převodníkové karty. Výstupy z převodníkové karty jsou připojeny k výkonovému zesilovači. Poháněný motor je stejnosměrný s permanentními magnety, tj. s konstantním buzením. Vinutí kotvy je připojeno k výstupu výkonového zesilovače.



obrázek IV.2-1: Stejnosměrný motor - schéma systému

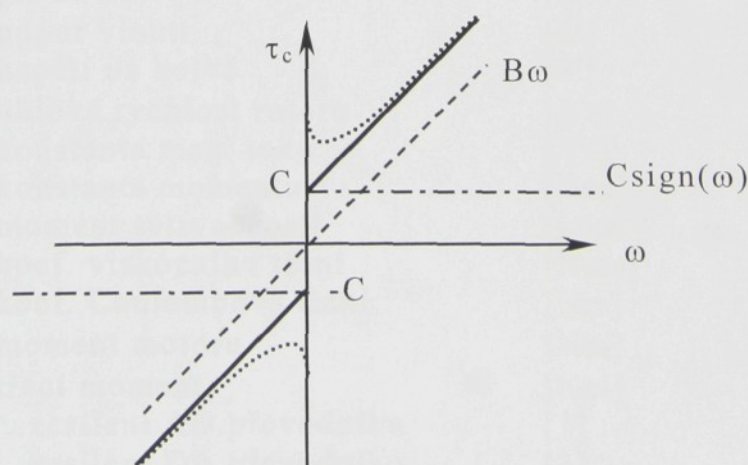
Výkonový zesilovač považujeme za ideální zdroj proudu řízený analogovým napěťovým signálem z digitálně-analogového

převodníku. Pro další popis jsme vytvořili principiální schéma elektromechanické části našeho systému (obrázek IV.2-2)



obrázek IV.2-2: Stejnosměrný motor - principiální schéma

Jak řídicí signál u , tak i výstupní signál y jsou interpretovány v počítači vždy tak, že jejich rozsah je v rozmezí $\langle -1,1 \rangle$. Převod mezi řídicí veličinou u , která je zadávána řídicím algoritmem počítače, a vstupním napětím výkonového zesilovače určuje konstanta digitálně-analogového převodníku k_{DA} . Výkonový zesilovač pak uvažujeme se zesílením k_z . Budíme-li obvod vinutí kotvy z ideálního zdroje proudu a indukčnost vinutí je malá, můžeme zanedbat ztráty napětí způsobené indukčností vinutí. R je odpor vinutí kotvy a I moment setrvačnosti rotoru, hřídele a připevněné řemenice.



obrázek IV.2-3: Coulombovo a viskózní tření

Uvažujme viskózní a Coulombovo tření v ložiskách. Brzdný moment je tedy nelineární funkcí úhlové rychlosti otáčení. Coulombovo a viskózní tření uvažujme ve zjednodušeném tvaru (viz. obrázek IV.2-3).

Při modelování vycházíme ze zjednodušeného modelu stejnosměrného motoru s konstantním buzením. Definujme:

$$i = \frac{1}{R}(u_k - k_b \varpi) \quad (\text{IV-7})$$

$$\tau = k_i i \quad (\text{IV-8})$$

$$\varpi = \frac{1}{k_{AD}} y \quad (\text{IV-9})$$

$$u_k = k_{DA} k_z u \quad (\text{IV-10})$$

$$I \dot{\omega} = \tau - \tau_c \quad (\text{IV-11})$$

po úpravě vztahů (IV-7) - (IV-11) dostaneme vztah (IV-12):

$$\dot{y} = \frac{k_i k_{DA} k_{AD} k_z}{RI} u - \frac{k_i k_b}{RI} y - C \text{sign}(y) - B y \quad (\text{IV-12})$$

a zjednodušeně:

$$\dot{y} = bu - ay - p \text{sign}(y), \quad (\text{IV-13})$$

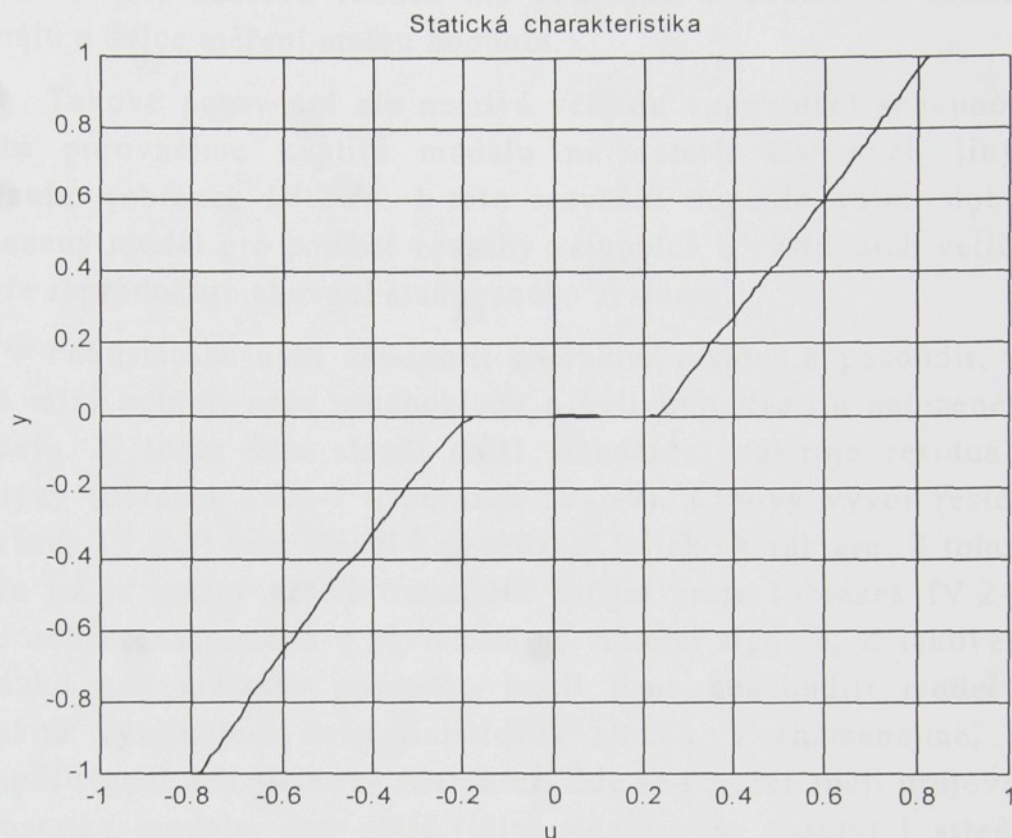
kde $a = \frac{k_i k_b}{RI} + B$, $b = \frac{k_i k_{DA} k_{AD} k_z}{RI}$, $p = C$,

u	řídící veličina	[V]
i	proud kotvou	[A]
R	odpor vinutí	[Ω]
u_k	napětí na kotvě	[V]
ω	úhlová rychlost rotoru	[1/s]
k_b	konstanta mag. toku	[Vs]
k_i	konstanta momentu	[Nm/A]
I	moment setrvačnosti	[kgm ²]
B	koef. viskózního tření	[Nms]
C	koef. Coulombova tření	[Nm]
τ	moment motoru	[Nm]
τ_c	třecí moment	[Nm]
k_{AD}	... zesílení AD převodníku	[1]
k_{DA}	... zesílení DA převodníku	[1]

a, b, p jsou hledané parametry modelovaného procesu

Při následné identifikaci považujeme dynamiku snímačů a převodníků za zanedbatelnou vzhledem k dynamice měřeného systému.

Podívejme se nejprve na naměřenou statickou charakteristiku systému (obrázek IV.2-4):



obrázek IV.2-4: Statická charakteristika stejnosměrného motorku

Z této naměřené statické charakteristiky (obrázek IV.2-4) je zřejmé (tak jak již naznačila matematicko-fyzikální analýza), že se jedná o nelineární systém, který lze snadno linearizovat v určitých oblastech. Jedná se o oblasti, ve kterých absolutní hodnota budícího signálu u překročí přibližně hodnotu 0,25. V těchto oblastech můžeme aplikovat metody identifikace lineárních systémů jako jsou integrační a derivační metoda. Tím získáme dynamiku naší identifikované soustavy. Pro potřeby identifikace lineárního modelu výše zmíněnými metodami jsme provedli měření v oblasti řídicího signálu většího než 0,25. Na takto získaných datech jsem již mohli použít nejprve *integrační metodu*. Data jsme ještě upravili tak, že úroveň řídicího signálu 0,3 jsme vzali jako za nulovou. Řád modelu

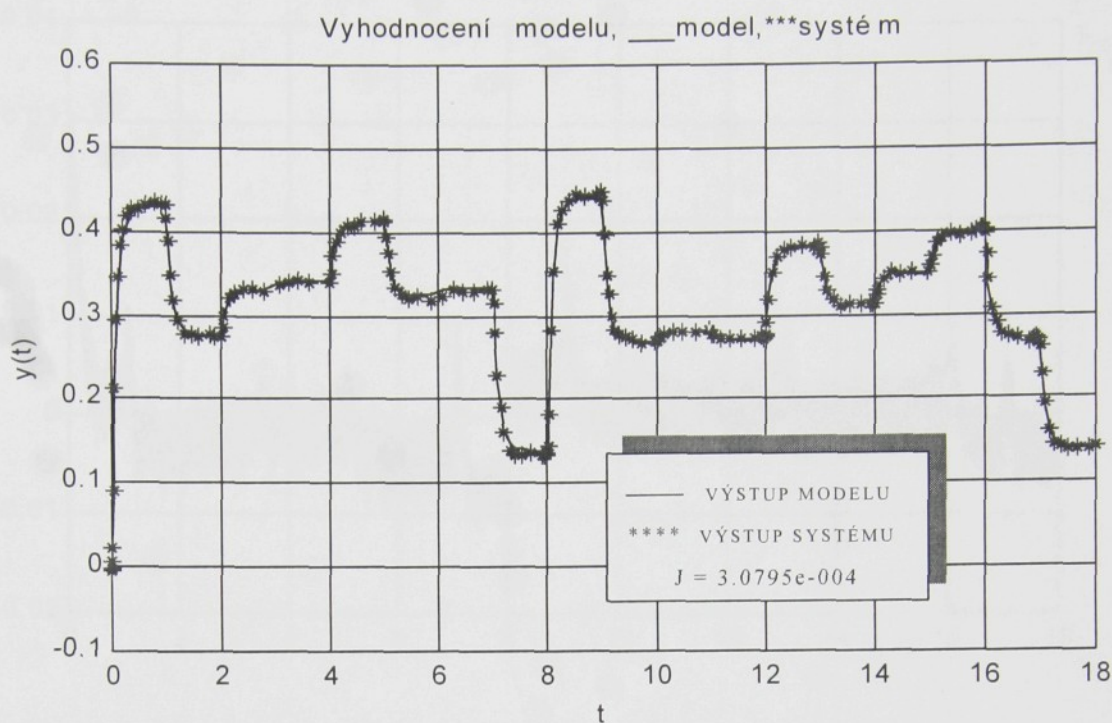
jsme zvolili 1, jak již vyplynulo z fyzikální analýzy systému. Po provedené optimalizaci dostáváme lineární model v této podobě:

$$\dot{y} = bu - ay, \text{ kde } a = 10.8668, b = 17.3704.$$

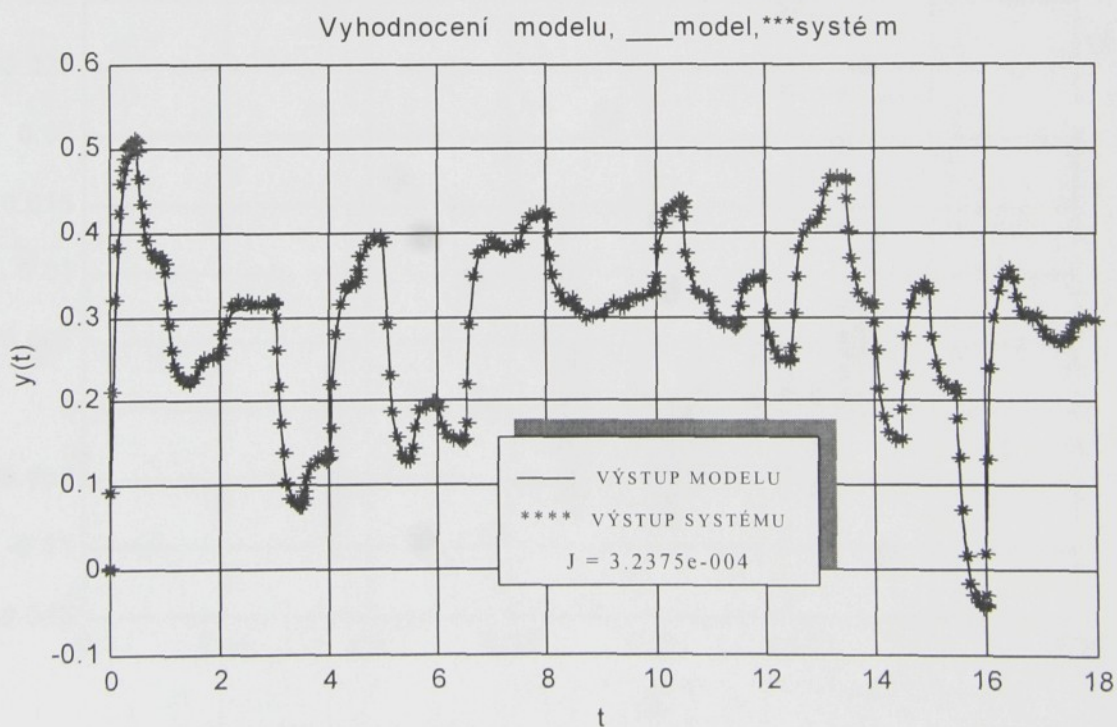
Podívejme se, jak si model vede v porovnání se skutečným systémem, a to nejprve na datech s pomocí kterých byl model vytvořen (obrázek IV.2-5). Výstupy modelu a systému jsou téměř identické a i účelová funkce má vzhledem k průměrné hodnotě signálu a délce měření malou hodnotu.

Takové porovnání ale nemívá velikou vypovídací schopnost, takže porovnejme kvalitu modelu na datech získaných jiným měřením (obrázek IV.2-6). I toto srovnání dopadlo velmi dobře. Nalezený model pro použité rozsahy vstupních a výstupních veličin dobře reprodukuje chování studovaného systému.

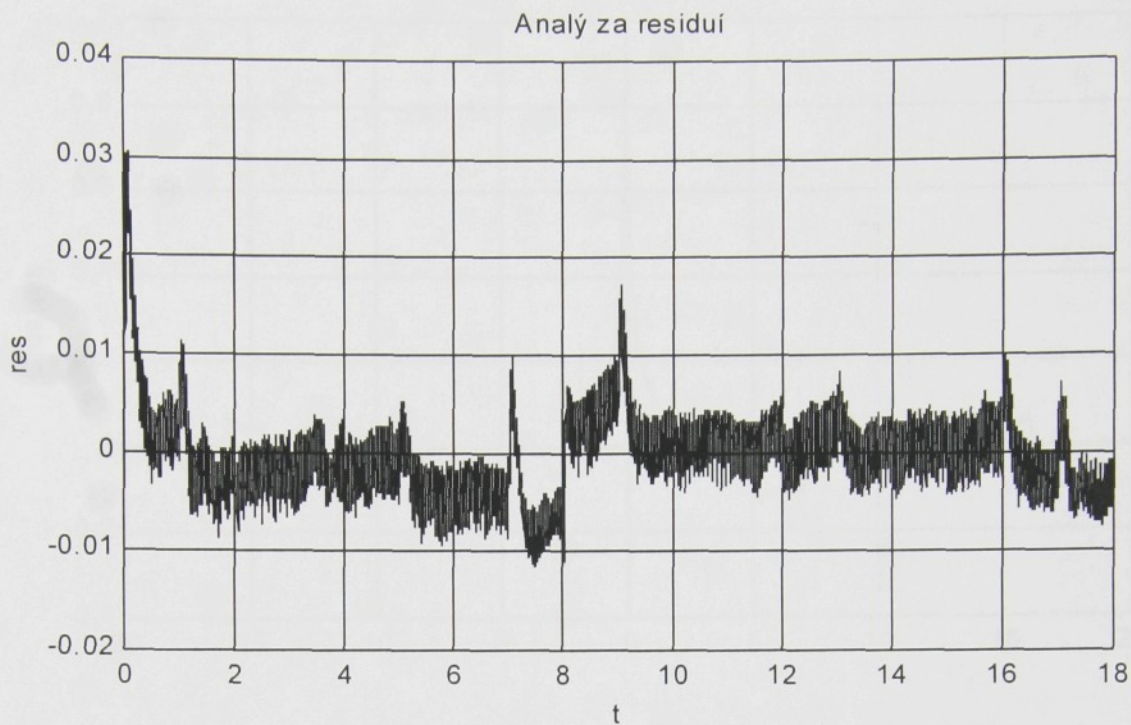
Pokusme se nyní zhodnotit charakter residuí a posoudit, do jaké míry splňují naše předpoklady a ovlivňují kvalitu nalezeného modelu. K tomu nám slouží další standardní nástroje residuální analýzy (obrázek IV.2-7 - obrázek IV.2-9). Časový vývoj residuí (obrázek IV.2-7) nám slouží k posouzení jejich charakteru. Z tohoto grafu již je patrný určitý trend. Na dalším grafu (obrázek IV.2-8) jsou zobrazena residua v závislosti na řídicím signálu. Z takového obrázku pak můžeme posoudit, jestli jsme neochudili model o nějakou významnou deterministickou složku. Poznamenejme, že vstupní signál měl skokový charakter. Zde se již zřetelněji projevují nedostatky modelu. Pro větší řídicí signál u se rozptýl i střední hodnota residuí zvyšuje. To může souviset s nelinearitou typu nasycení. A na závěr vyhodnoťme autokovarianční funkci residuí (obrázek IV.2-9). Ta nám potvrzuje, že její průběh nemá charakter autokov. funkce nekorelovaného šumu. Z těchto všech předvedených kritérií na jakost modelu vidíme, že model není pravdivým popisem systému (ani jím být nemůže), ale i když máme výhrady k tomuto modelu lze ho použít v oblasti automatického řízení. Při potřebě nalézt model přesnější je doporučeno použít některou z iteračních metod, jak bude ukázáno dále.



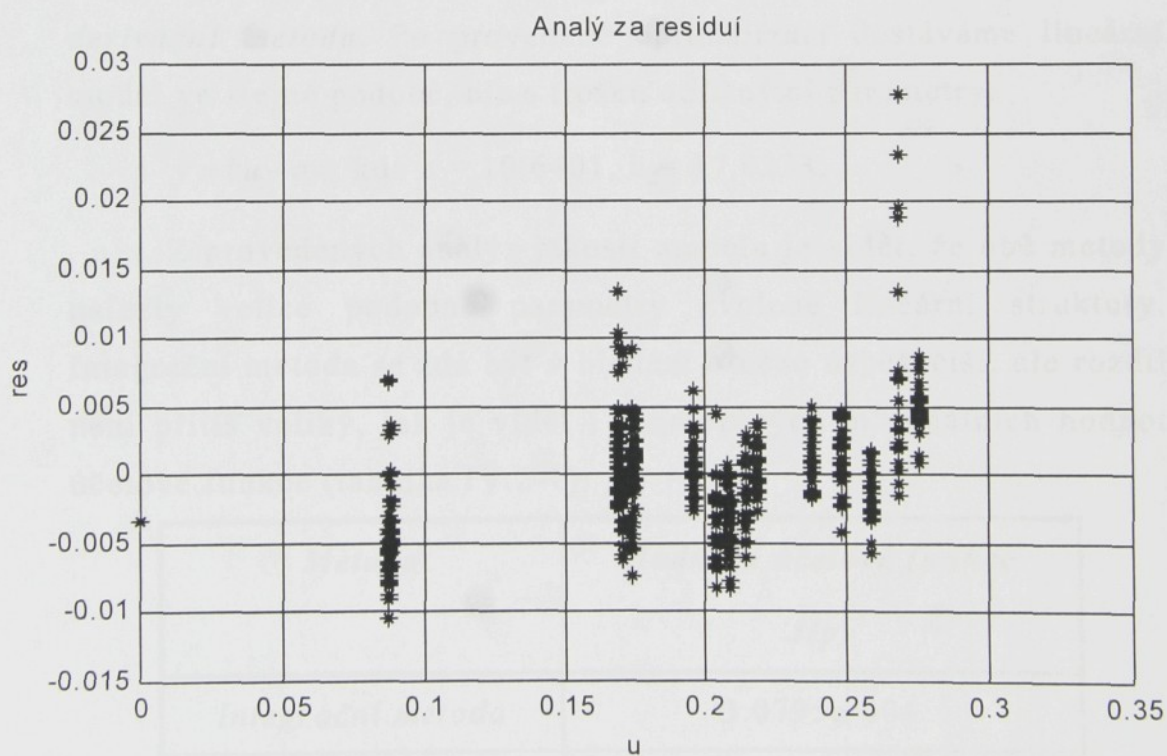
obrázek IV.2-5: Porovnání výstupu modelu a systému ss motorku na datech použitých pro optimalizaci



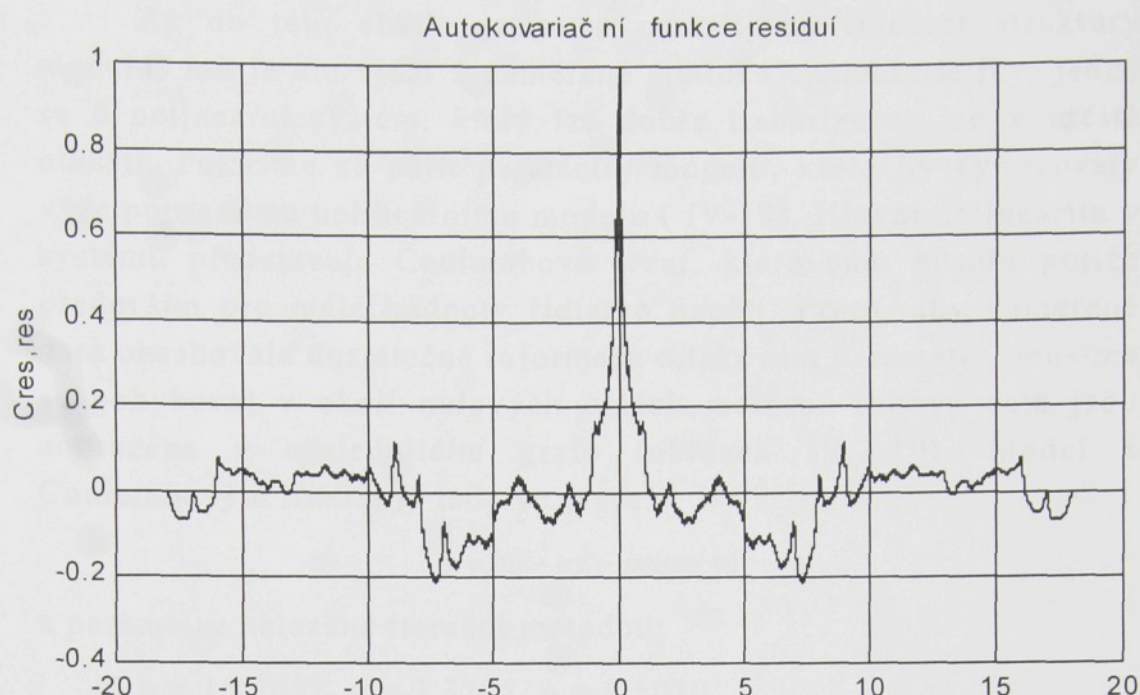
obrázek IV.2-6: Porovnání výstupu modelu a systému ss motorku na jiných datech než byla použita pro optimalizaci



obrázek IV.2-7: Časová posloupnost residuí modelu ss motorku



obrázek IV.2-8: Graf residuí v závislosti na budícím signálu



obrázek IV.2-9: Autokovarianční funkce residuí modelu ss motorku

Nyní se podívejme, jak si bude se stejnými daty vést *derivační metoda*. Po provedené optimalizaci dostáváme lineární model ve stejné podobě, ale s trošku odlišnými parametry:

$$\dot{y} = bu - ay, \text{ kde } a = 10.6401, b = 17.0223.$$

Z provedených analýz jakosti modelu je vidět, že obě metody našly velice podobné parametry zvolené lineární struktury. Integrační metoda se zdá být v hledání o něco úspěšnější, ale rozdíl není příliš veliký, jak je vidět i z nalezených minimálních hodnot účelové funkce (tabulka IV.2-1):

<i>Metoda</i>	<i>Hodnota účelové funkce</i> <i>J(p)</i>
<i>Integrační metoda</i>	3.0795e-004
<i>Derivační metoda</i>	3.3151e-004

tabulka IV.2-1: Hodnoty účelové funkce pro systém ss motorku

Až do této chvíle jsme se omezili na lineární struktury modelu. Jak je ale vidět z naměřené statické charakteristiky, jedná se o nelineární systém, který lze dobře linearizovat jen v určité oblasti. Pokusme se najít parametry modelu, které by vyhovovaly výše popsanému nelineárnímu modelu (IV-13). Hlavní nelinearita v systému představuje Coulombovo tření, které nám působí potíže především pro malé hodnoty řídicího napětí. Proto, aby naměřená data obsahovala dostatečné informace o takovém parametru, musíme se pohybovat v okolí nulových otáček motoru. Taková data jsou zobrazena v následujícím grafu (obrázek IV.2-10). Model s Coulombovým třením je tedy ve tvaru:

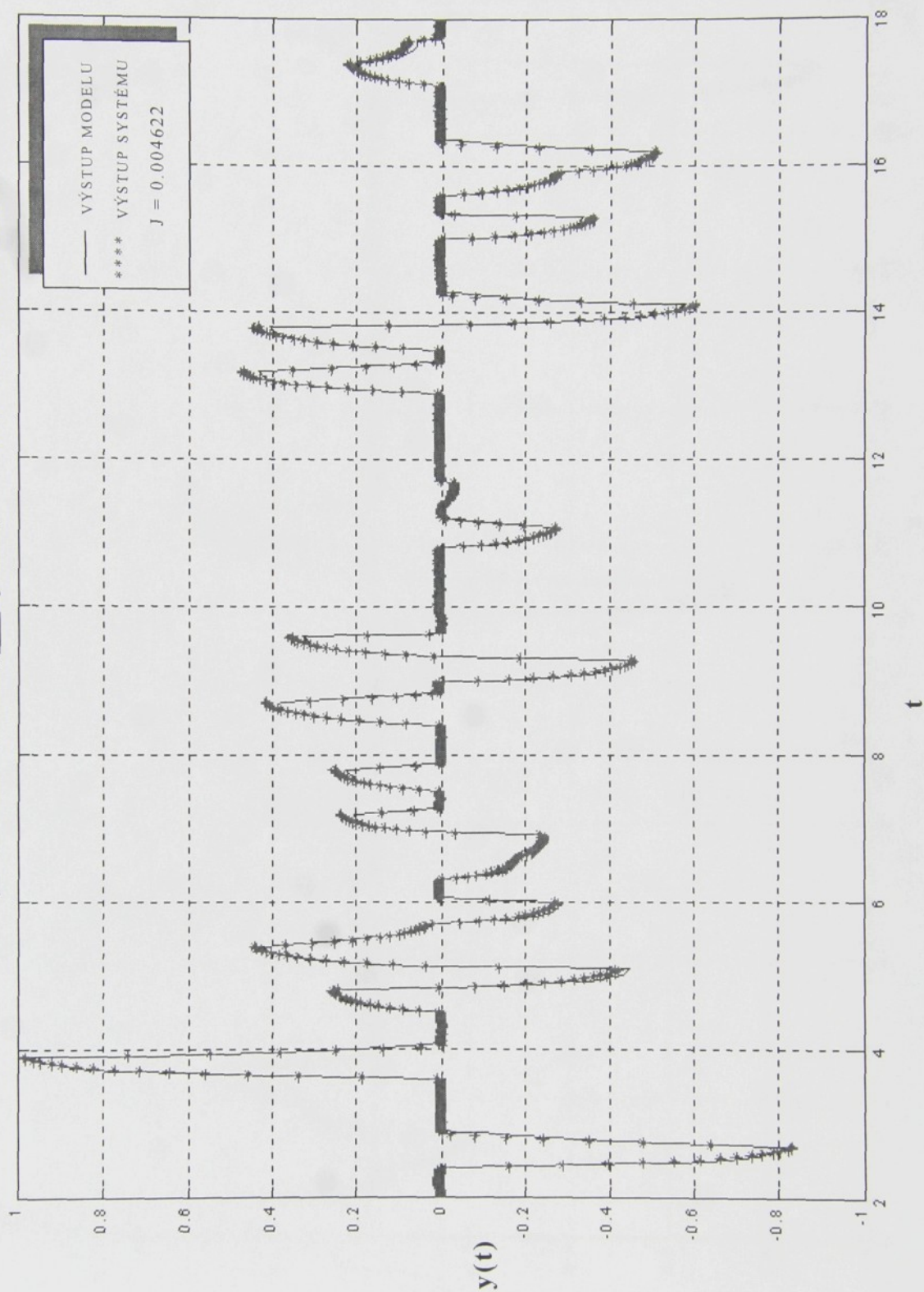
$$\dot{y} = bu - ay - p \operatorname{sign}(y)$$

a parametry nalezené iterační metodou:

$$b = 15.7022, a = 3.3295, p = 9.5030.$$

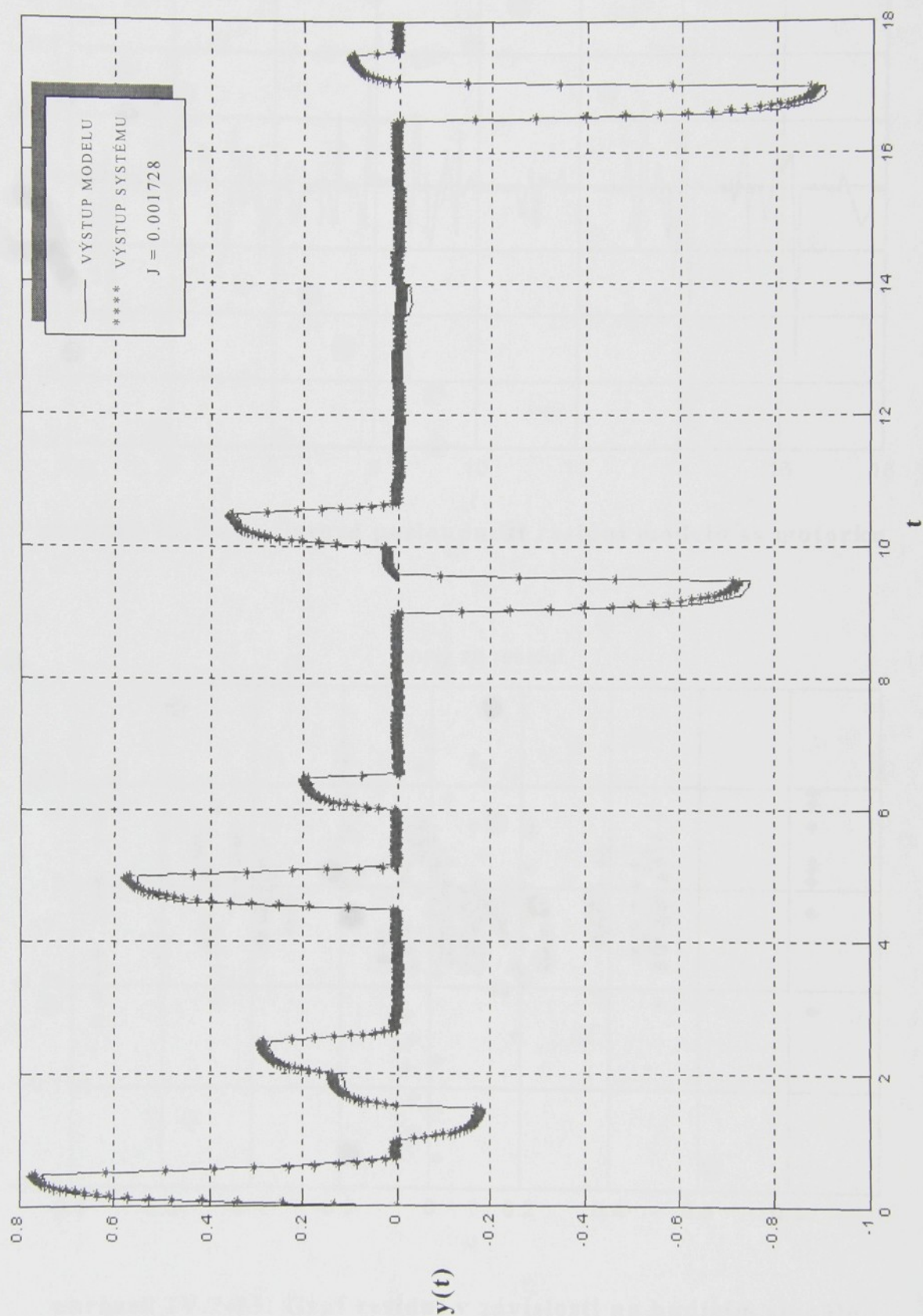
Podívejme se na porovnání výstupu nelineárního modelu a vyšetřovaného systému ss motorku (obrázek IV.2-10). Pro toto srovnání byla použita stejná data jako pro optimalizaci. Výstupy modelu a systému jsou velmi podobné a i účelová funkce má poměrně malou hodnotu. Na dalším obrázku (obrázek IV.2-11) je zobrazeno srovnání reakcí modelu a systému výměníku na jiný signál, než byl použit pro optimalizaci. I toto srovnání dopadlo velmi dobře. Nalezený model pro použité rozsahy vstupních a výstupních veličin dobře reprodukuje chování studovaného systému.

Vyhodnocení modelu, __Systém, ** Model s Coulom. třením

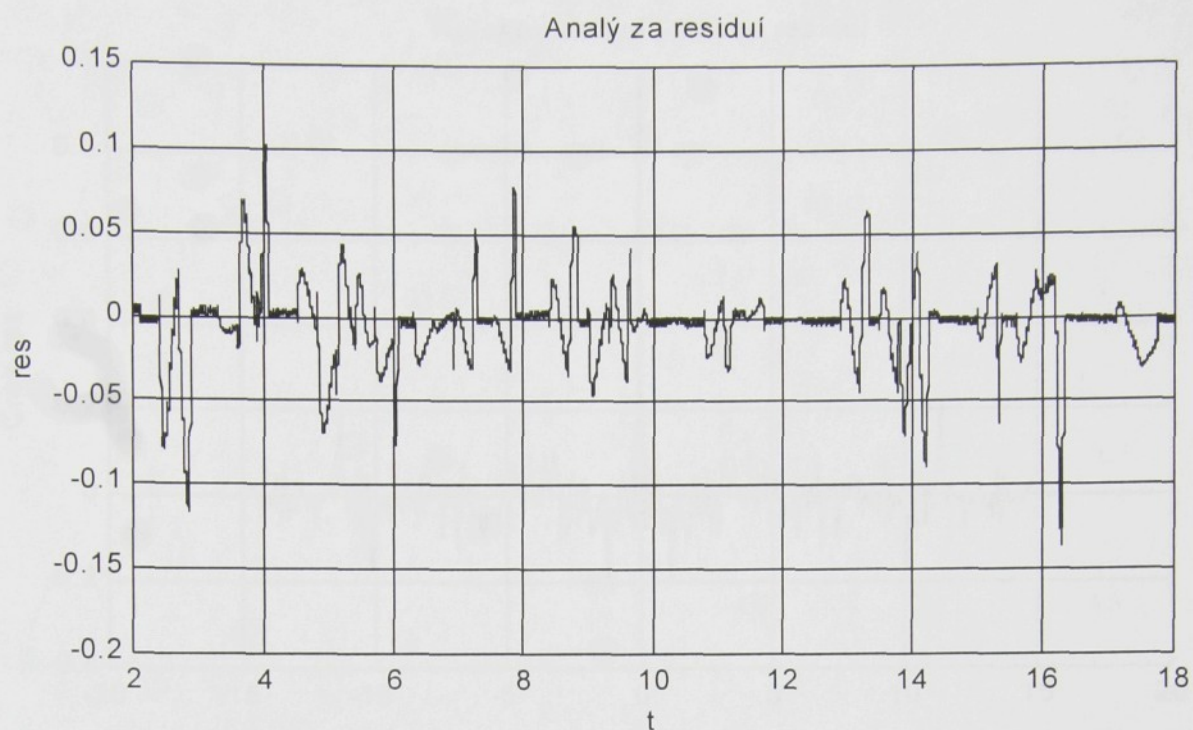


obrázek IV.2-10: Porovnání výstupu modelu a systému ss
motorku na datech použitých pro optimalizaci

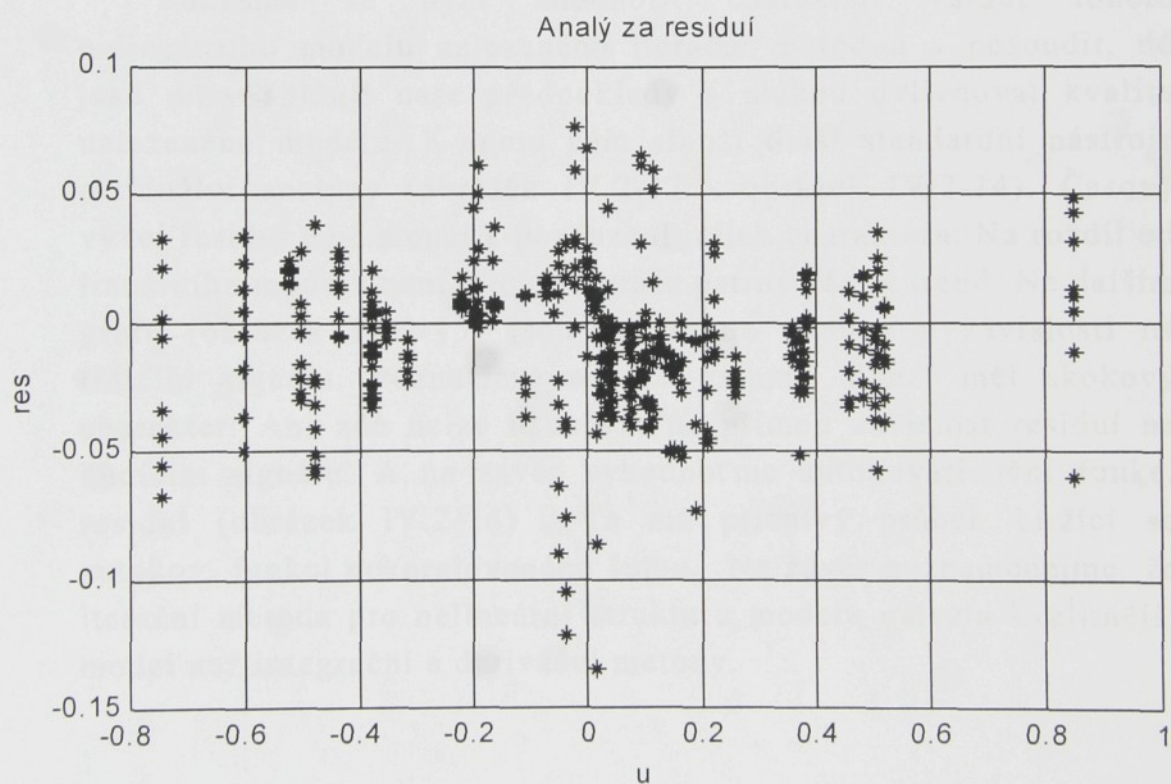
Vyhodnocení modelu. Systém.*** Model s Coulom. třením



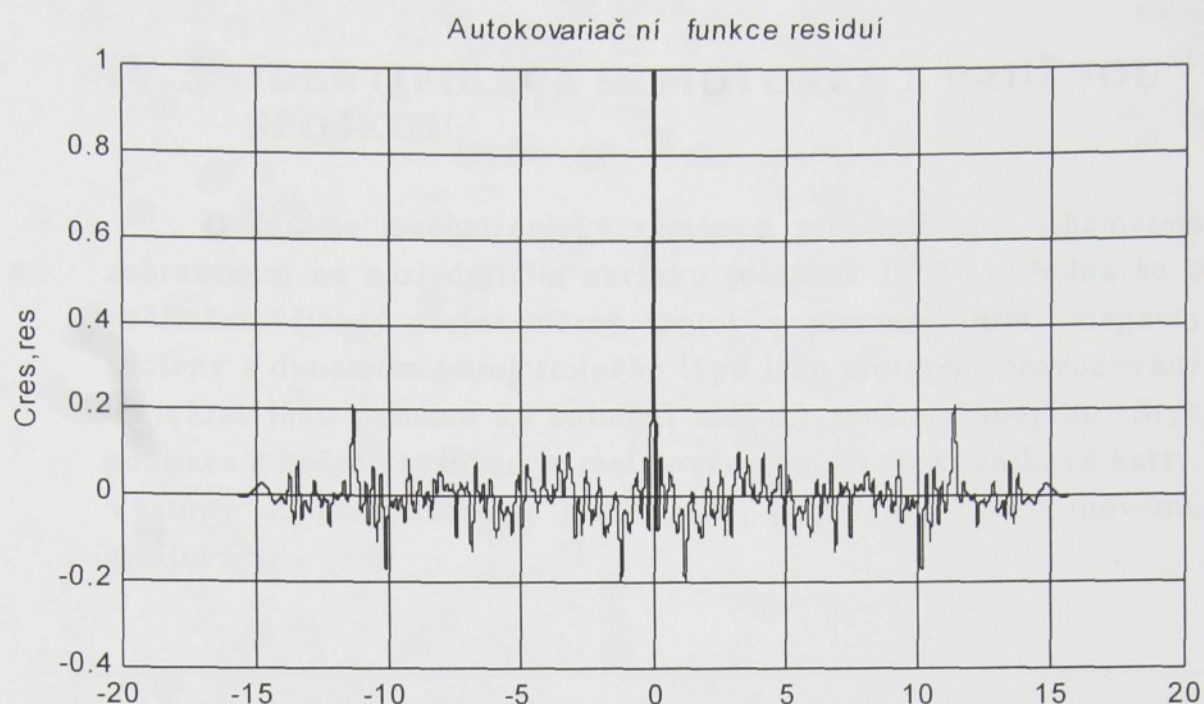
obrázek IV.2-11: Porovnání výstupu modelu a systému ss
motorku na jiných datech než byla použita pro optimalizaci



obrázek IV.2-12: Časová posloupnost residuí modelu ss motorku



obrázek IV.2-13: Graf residuí v závislosti na budícím signálu

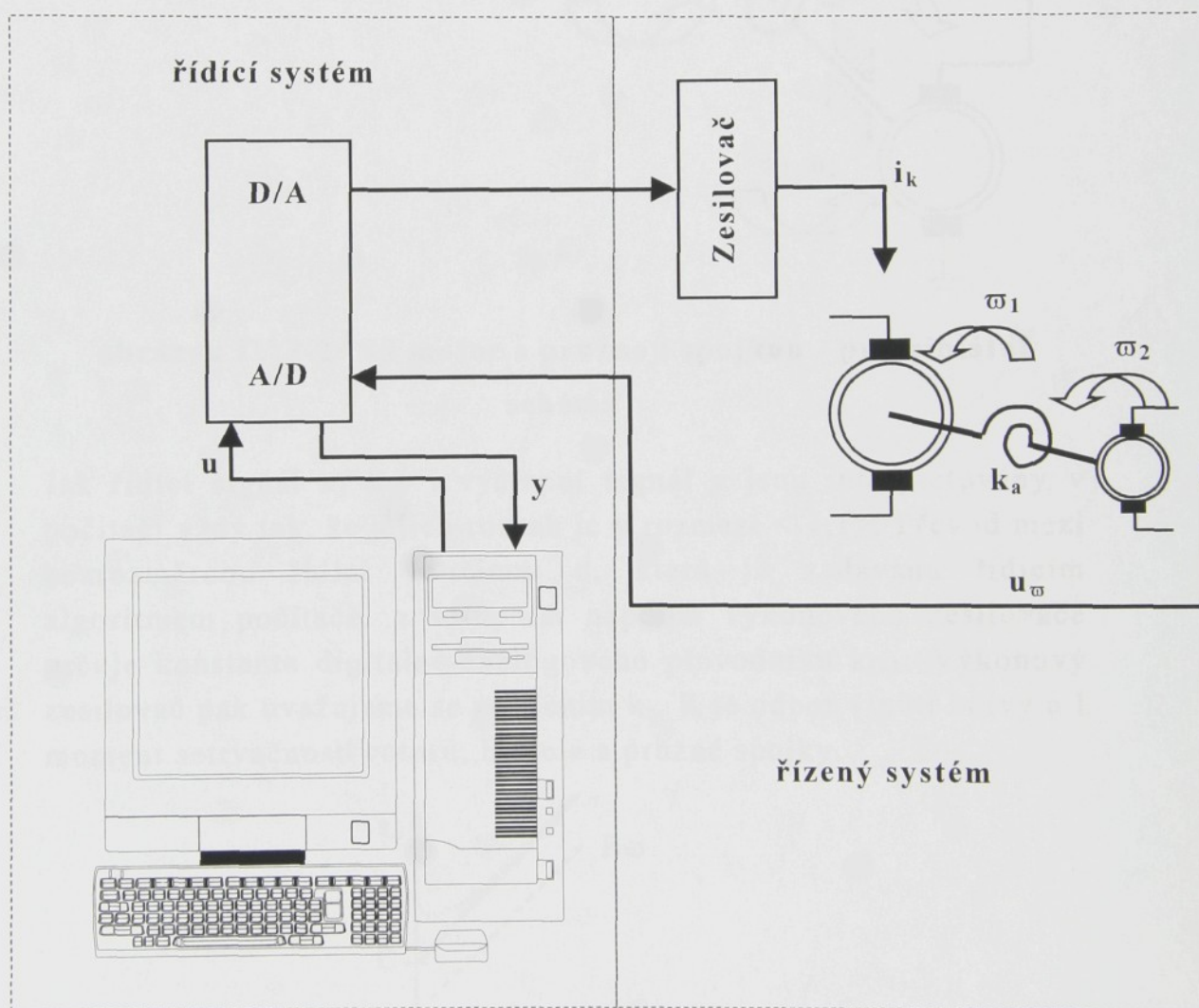


obrázek IV.2-14: Autokovarianční funkce residuí

Pokusme se nyní zhodnotit charakter residuí tohoto nelineárního modelu nalezeného iterační metodou a posoudit, do jaké míry splňují naše předpoklady a mohou ovlivňovat kvalitu nalezeného modelu. K tomu nám slouží další standardní nástroje residuální analýzy (obrázek IV.2-12 - obrázek IV.2-14). Časový vývoj residuí nám slouží k posouzení jejich charakteru. Na rozdíl od lineárního modelu není z tohoto grafu patrný žádný trend. Na dalším grafu (obrázek IV.2-13) jsou zobrazena residua v závislosti na řídicím signálu. Poznamenejme, že vstupní signál měl skokový charakter. Ani zde nelze usuzovat na přímou závislost residuí na budícím signálu. A na závěr vyhodnotíme autokovarianční funkci residuí (obrázek IV.2-14). Ta má příznivý průběh blíží se autokov. funkci nekorelovaného šumu. Na závěr poznamenejme, že iterační metoda pro nelineární strukturu modelu našla kvalitnější model než integrační a derivační metody.

IV.3 IDENTIFIKACE SS MOTORKU S PRUŽNOU SPOJKOU

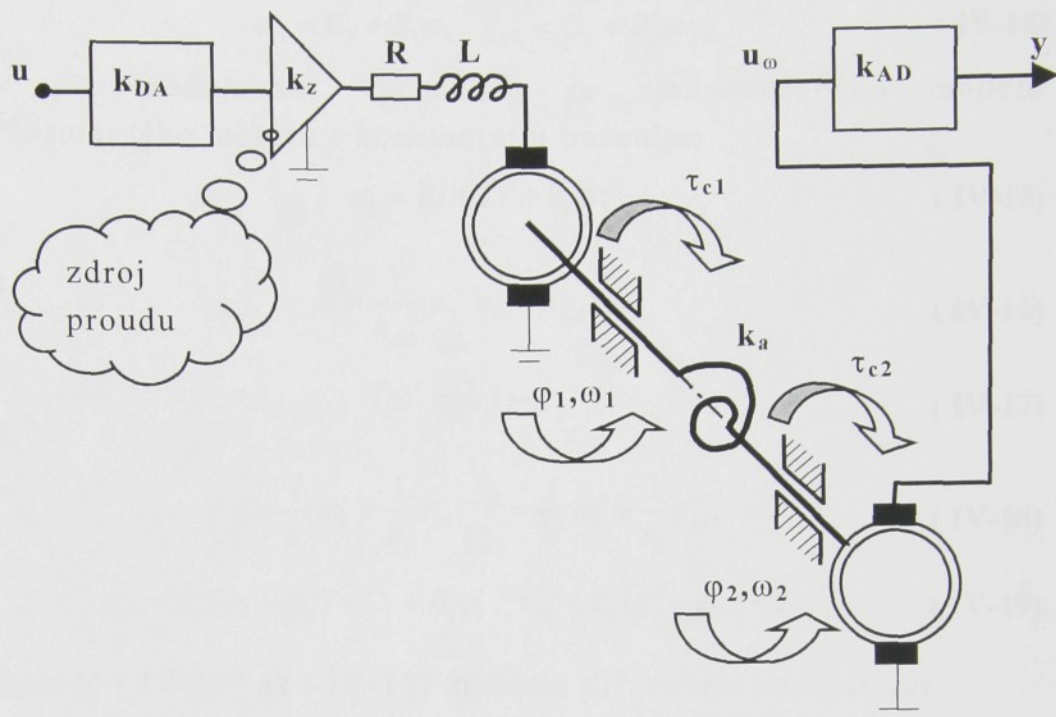
Uvažujme mechatronický systém s principiálním schématem zobrazeném na následujícím obrázku (obrázek IV.3-1). Jedná se o počítačem řízený stejnosměrný motor s permanentními magnety spojený s dynamem (stroj stejného typu jako motorek, provozovaný tentokrát jako dynamo ke snímání otáček) pružnou spojkou. Styk počítače a našeho systému je realizován pomocí převodníkové karty. Výstupy z převodníkové karty jsou připojeny k výkonovému zesilovači.



obrázek IV.3-1: SS motorek s pružnou spojkou - schéma systému

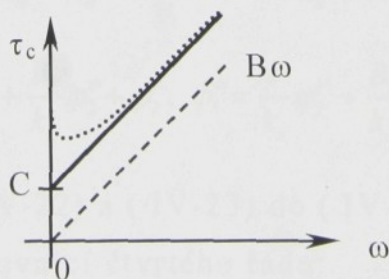
Poháněný motor je stejnosměrný s permanentními magnety, tj. s konstantním buzením. Vinutí kotvy je připojeno k výstupu

výkonového zesilovače. Výkonový zesilovač je řízený analogovým napětovým signálem z digitálně-analogového převodníku. Pro další popis jsme vytvořili principiální schéma elektromechanické části našeho systému (obrázek IV.3-2):



obrázek IV.3-2: SS motor s pružnou spojkou - principiální schéma

Jak řídicí signál u , tak i výstupní signál y jsou interpretovány v počítači vždy tak, že jejich rozsah je v rozmezí $\langle -1, 1 \rangle$. Převod mezi bezrozměrnou řídicí veličinou u , která je zadávána řídicím algoritmem počítače, a vstupním napětím výkonového zesilovače určuje konstanta digitálně-analogového převodníku k_{DA} . Výkonový zesilovač pak uvažujeme se zesílením k_z . R je odpor vinutí kotvy a L moment setrvačnosti rotoru, hřídele a pružné spojky.



obrázek IV.3-3: Coulombovo a viskózní tření

Uvažujme viskózní a Coulombovo tření v ložiskách. Brzdný moment je tedy nelineární funkcí úhlové rychlosti otáčení. Coulombovo a viskózní tření uvažujme ve zjednodušeném tvaru (obrázek IV.3-3) s přihlédnutím k tomu, že se systém otáčí pouze v jednom směru (ten uvažujme jako kladný):

$$\tau_{c1} = C_1 + B_1 \varpi_1, \quad \tau_{c2} = C_2 + B_2 \varpi_2, \quad (\text{IV-14})$$

dále při modelování vycházíme ze zjednodušeného modelu stejnosměrného motoru s konstantním buzením:

$$u_k = Ri + Li' + k_b \omega, \quad (\text{IV-15})$$

$$\omega = \frac{1}{k_{AD}} y, \quad u_k = k_{DA} k_z u, \quad (\text{IV-16})$$

$$I \omega' = k_i i - \tau_{z1}, \quad (\text{IV-17})$$

$$i = \frac{I}{k_i} \omega_1' + \frac{1}{k_i} \tau_{z1}, \quad i' = \frac{I}{k_i} \omega_1'' + \frac{1}{k_i} \tau_{z1}', \quad (\text{IV-18})$$

$$\tau_{z1} = k_a (\varphi_1 - \varphi_2) + C_1 + B_1 \varphi_1', \quad \tau_{z1}' = k_a (\varphi_1' - \varphi_2') + B_1 \varphi_1'', \quad (\text{IV-19})$$

po úpravě (IV-15) až (IV-18) získáme dif. rovnici popisující dynamiku otáčení rotoru motoru (IV-20):

$$\frac{LI}{k_i} \omega_1'' + \frac{RI}{k_i} \omega_1' + k_b \omega_1 = k_{DA} k_z u - \frac{R}{k_i} \tau_{z1} - \frac{L}{k_i} \tau_{z1}' \quad (\text{IV-20})$$

a na straně zátěže:

$$I \varphi_2'' = k_a (\varphi_1 - \varphi_2) - \tau_{c2} = k_a (\varphi_1 - \varphi_2) - C_2 - B_2 \varphi_2'. \quad (\text{IV-21})$$

Vyjádříme z rovnice (IV-21) polohu hřídele dynamy:

$$\varphi_1 = \frac{I}{k_a} \varphi_2'' + \frac{C_2}{k_a} + \frac{B_2}{k_a} \varphi_2' + \varphi_2, \quad \varphi_1' = \frac{I}{k_a} \varphi_2''' + \frac{B_2}{k_a} \varphi_2'' + \varphi_2', \quad (\text{IV-22})$$

$$\varphi_1'' = \frac{I}{k_a} \varphi_2^{IV} + \frac{B_2}{k_a} \varphi_2''' + \varphi_2'', \quad \varphi_1''' = \frac{I}{k_a} \varphi_2^V + \frac{B_2}{k_a} \varphi_2^{IV} + \varphi_2'''. \quad (\text{IV-23})$$

Po dosazení (IV-22) a (IV-23) do (IV-20) a (IV-19) a úpravě získáme dif. rovnici čtvrtého řádu:

$$y^{IV} + a_3 y''' + a_2 y'' + a_1 y' + a_0 y + D = b_0 u, \quad (\text{IV-24})$$

kde jednotlivé koeficienty jsou:

$$a_3 = \frac{1}{I}(B_1 + B_2) + \frac{R}{L}, \quad a_2 = \frac{R}{LI}(B_1 + B_2) + \frac{k_b k_i}{LI} + \frac{1}{I^2} B_1 B_2 + \frac{2k_a}{I}, \quad (\text{IV-25})$$

$$a_1 = \frac{k_a}{I^2}(B_1 + B_2) + \frac{2k_a R}{LI} + \frac{B_2}{LI^2}(k_b k_i + RB_1), \quad (\text{IV-26})$$

$$a_0 = \frac{k_a}{LI^2}(k_i k_b + RB_2 + RB_1), \quad (\text{IV-27})$$

$$D = \frac{k_a R^2 C_2 C_1}{LI^2 k_i}, \quad b_0 = \frac{k_{DA} k_z k_i k_a}{LI^2}, \quad (\text{IV-28})$$

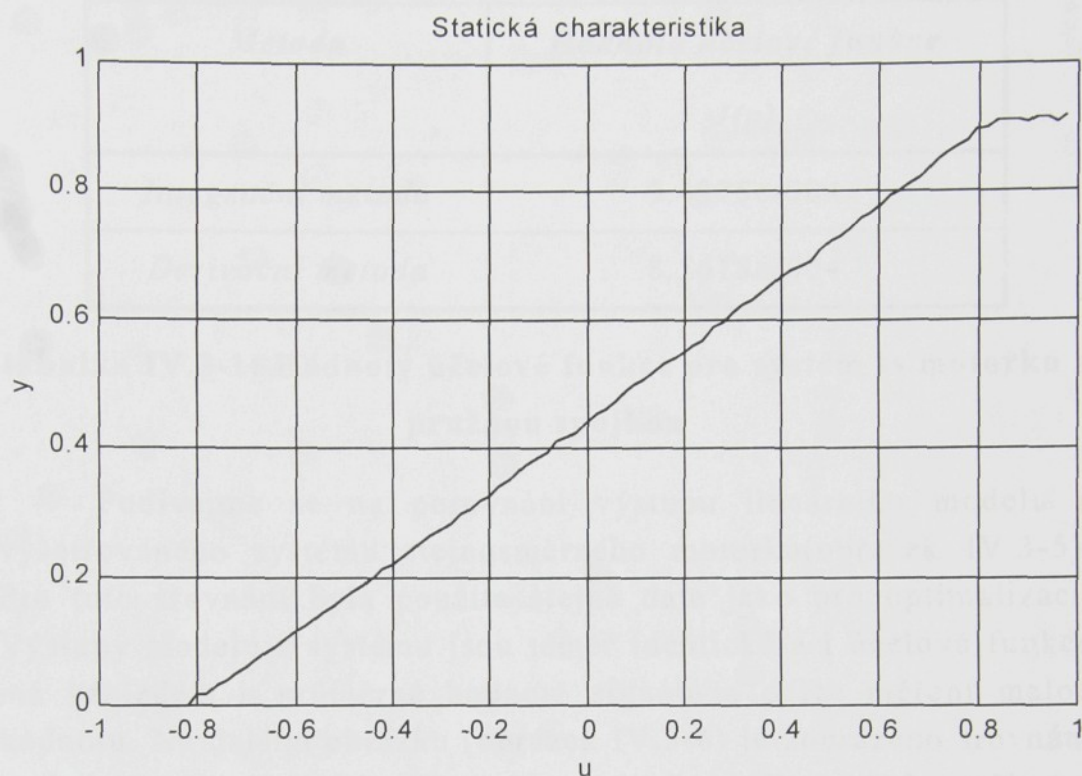
kde:

u	řídící veličina	[V]
i	proud kotvou	[A]
R	odpor vinutí	[Ω]
L	indukčnost vinutí	[H]
u _k	napětí na kotvě	[V]
ω	úhlová rychlost rotoru	[1/s]
φ	pootočení rotoru	[1]
k _b	konstanta mag. toku	[Vs]
k _i	konstanta momentu	[Nm/A]
k _a	konstanta pružiny	[Nm ⁻¹]
I	moment setrvačnosti	[kgm ²]
B	koef. viskózního tření	[Nms]
C	koef. Coulombova tření	[Nm]
τ	moment motoru	[Nm]
τ _c	třecí moment	[Nm]
τ _{z1}	zátěžový moment motoru	[Nm]
k _{AD} ...	zesílení AD převodníku	[1]
k _{DA} ...	zesílení DA převodníku	[1]

Podívejme se nejprve na naměřenou statickou charakteristiku systému (obrázek IV.3-4).

Naměřená statická charakteristika systému odpovídá výše uvedeným předpokladům a fyzikální analýze (model s Coulombovým třením). Systém lze tedy považovat za lineární v oblastech, ve kterých hodnota řídícího signálu překročí přibližně hodnotu -0,8. V těchto oblastech můžeme aplikovat metody identifikace lineárních systémů (za předpokladu, že přesuneme nulovou hodnotu budícího signálu modelu do lineární oblasti) jako je integrační a derivační metoda. Tím získáme dynamiku studované identifikované soustavy. Pro potřeby identifikace lineárního modelu

výše zmíněnými metodami jsme provedli měření splňující podmínku řídicího signálu na linearitu systému, tzn. při měření nám signál neklesl pod hodnotu -0,8.



obrázek IV.3-4: Statická charakteristika ss motorku s pružnou spojkou

I když z matematicko-fyzikální analýzy vyplývá přenosová funkce se čtyřmi póly a žádnou nulou, pro integrační a derivační metody nejlépe fungoval model se čtyřmi póly a jednou nulou:

$$y^{IV} + a_3 y''' + a_2 y'' + a_1 y' + a_0 y = b_0 u + b_1 u', \quad (\text{IV-29})$$

kde nalezené parametry integrační metodou jsou:

$$a_3 = 15.4914, a_2 = 47.3217, a_1 = 280.2342, a_0 = 163.8359$$

$$b_1 = 155.3839, b_0 = 94.2985.$$

Nalezené parametry derivační metodou pro (IV-29) jsou:

$$a_3 = 15.8499, a_2 = 47.8792, a_1 = 281.5456, a_0 = 163.9327$$

$$b_1 = 155.5646, b_0 = 94.3542.$$

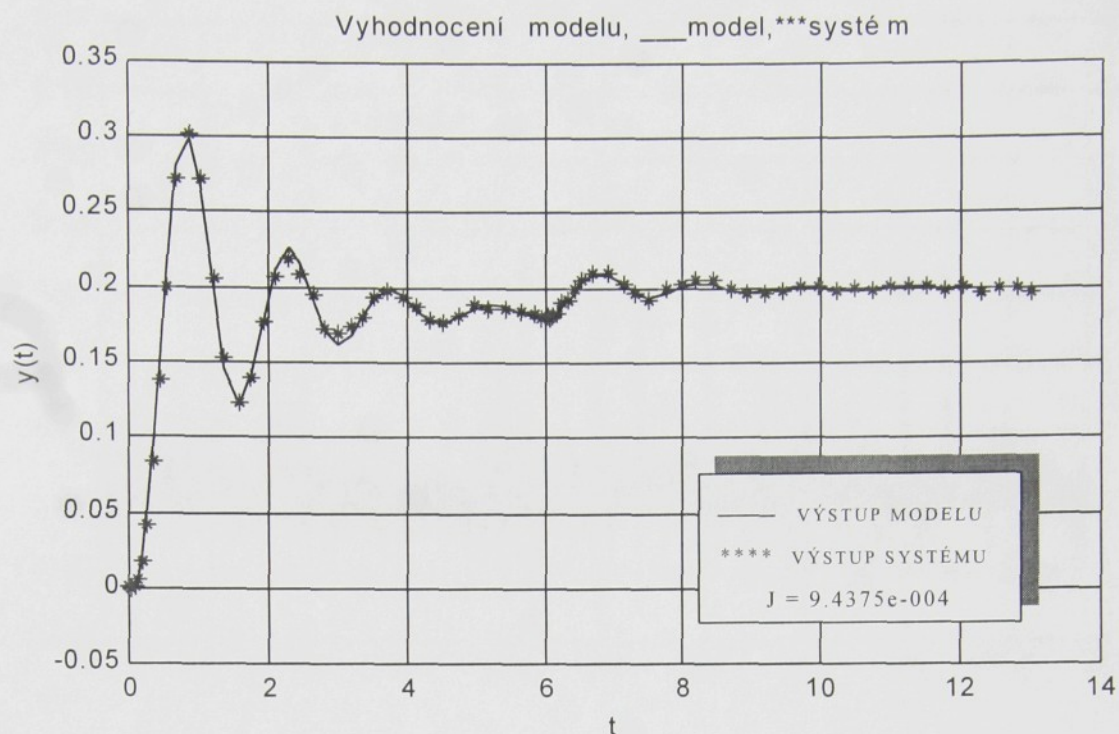
Obě metody (integrační a derivační) našly velice podobné parametry zvolené lineární struktury. Derivační metoda byla v

hledání o něco úspěšnější, ale rozdíl není příliš veliký, jak je i vidět z nalezených minimálních hodnot účelové funkce (viz. tabulka IV.3-1):

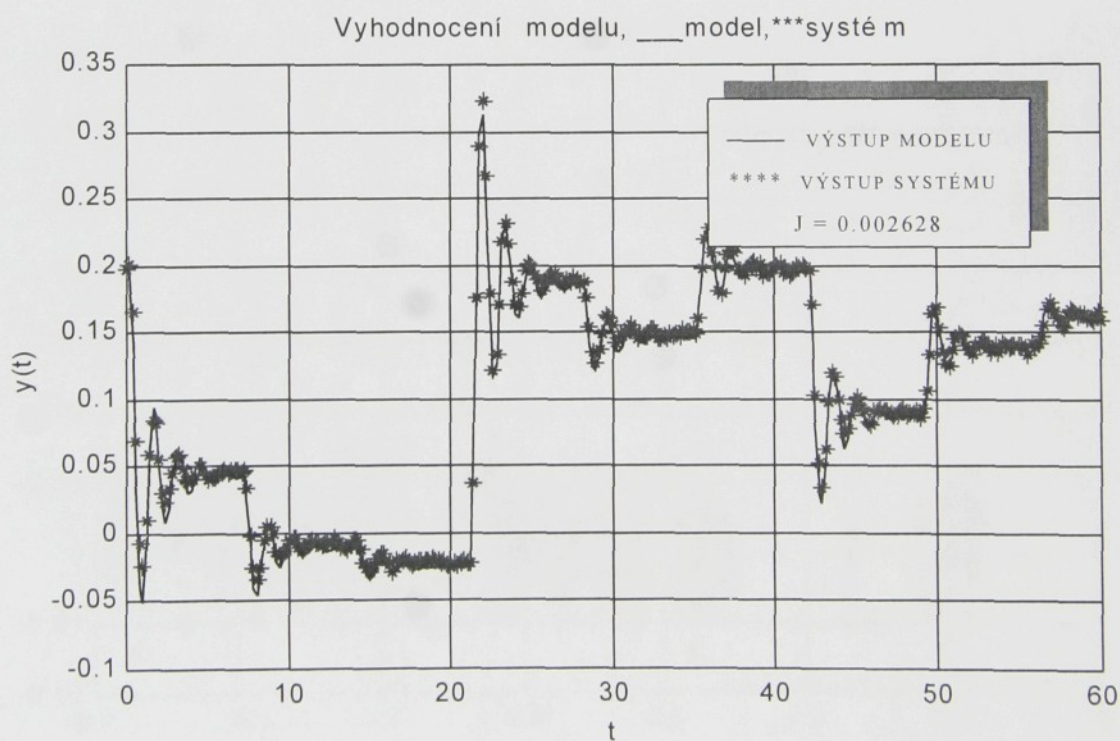
<i>Metoda</i>	<i>Hodnota účelové funkce</i> <i>J(p)</i>
<i>Integrační metoda</i>	9.4375e-004
<i>Derivační metoda</i>	8.5618e-004

tabulka IV.3-1: Hodnoty účelové funkce pro systém ss motorku s pružnou spojkou

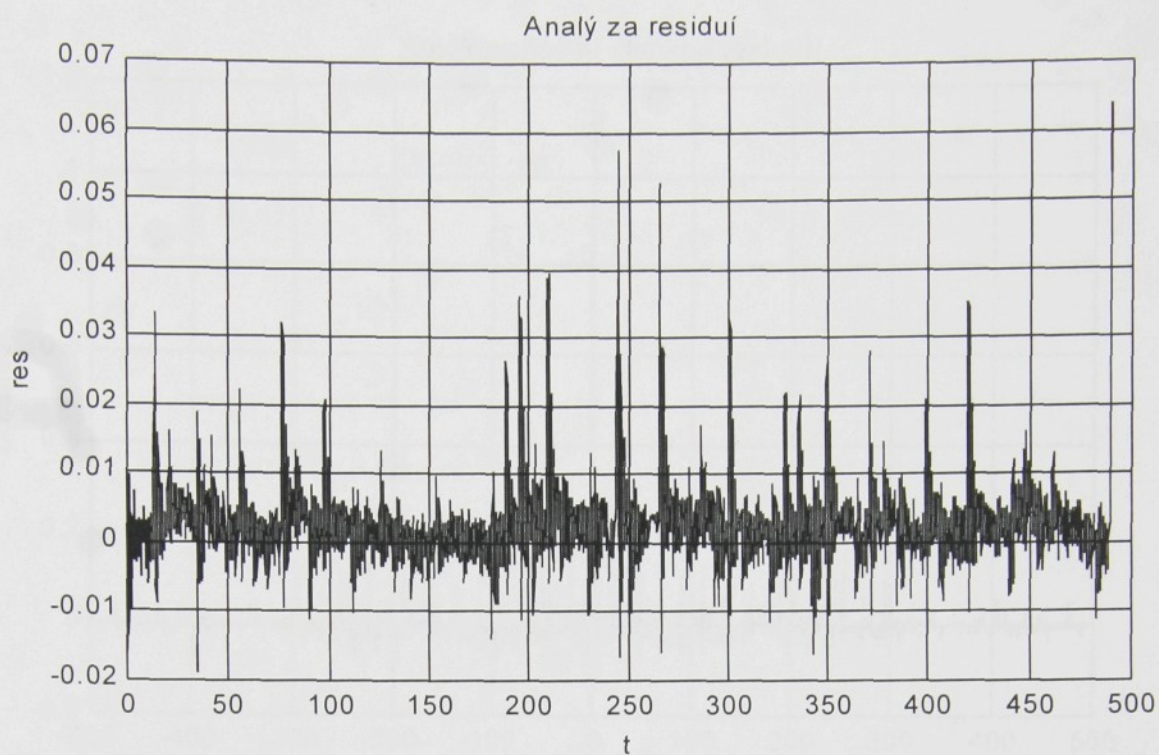
Podívejme se na porovnání výstupu lineárního modelu a vyšetřovaného systému stejnosměrného motorku (obrázek IV.3-5). Pro toto srovnání byla použita stejná data jako pro optimalizaci. Výstupy modelu a systému jsou téměř identické a i účelová funkce má vzhledem k průměrné hodnotě signálu a délce měření malou hodnotu. Na dalším obrázku (obrázek IV.3-6) je zobrazeno srovnání reakcí modelu a systému výměníku na jiný signál než byl použit pro optimalizaci. I toto srovnání dopadlo velmi dobře. Nalezený model pro použité rozsahy vstupních a výstupních veličin dobře reprodukuje chování studovaného systému.



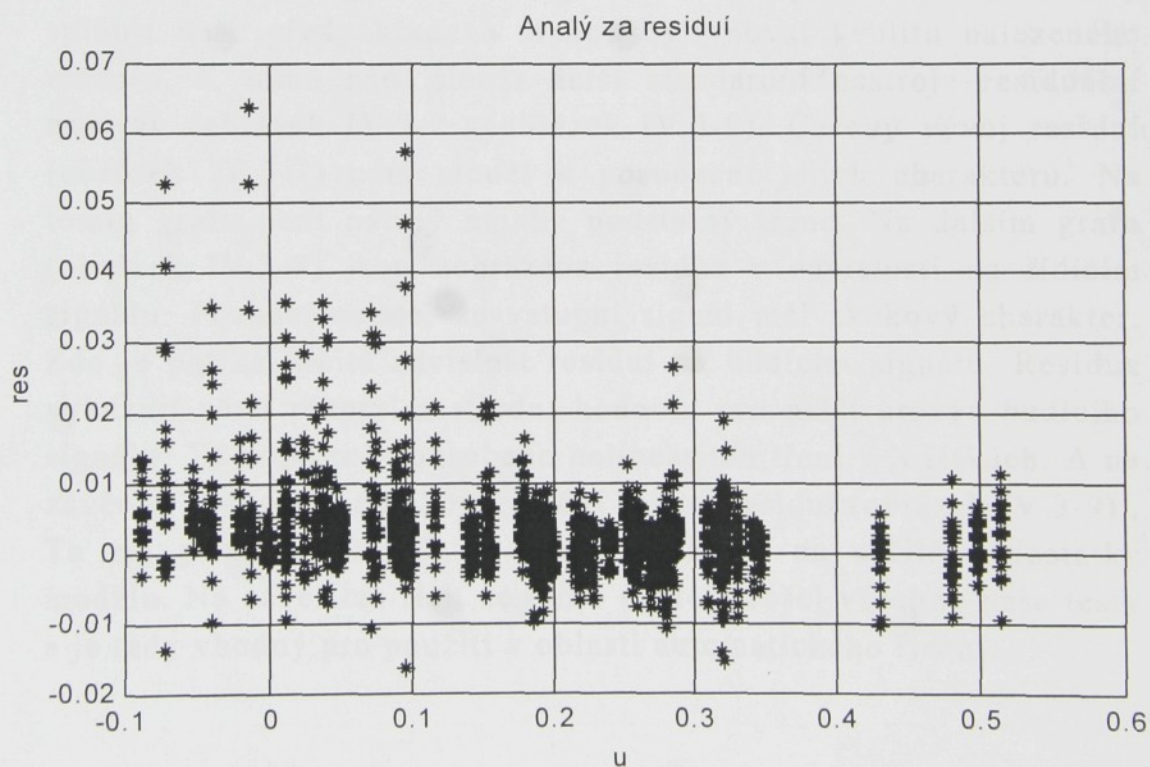
obrázek IV.3-5: Porovnání výstupu modelu a systému ss motorku s pružnou spojkou na datech použitých pro optimalizaci



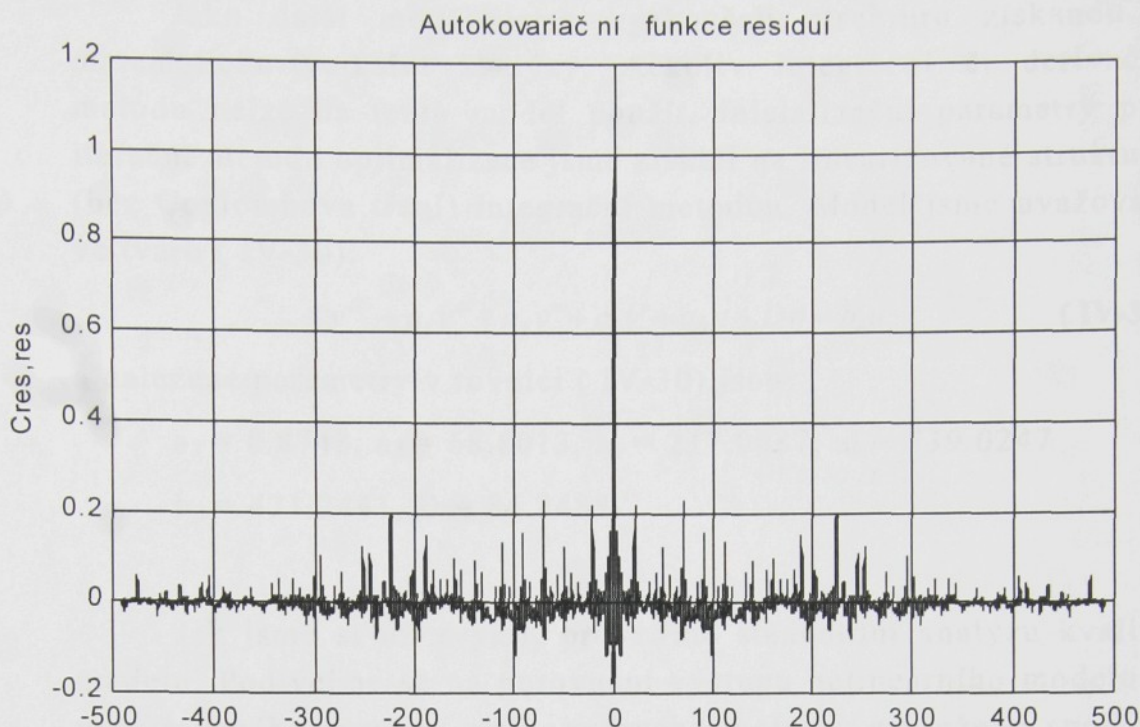
obrázek IV.3-6: Porovnání výstupu modelu a systému ss motorku s pružnou spojkou na jiných datech než byla použita pro optimalizaci



obrázek IV.3-7: Časová posloupnost residuí modelu ss motorku s pružnou spojkou



obrázek IV.3-8: Graf residuí v závislosti na budícím signálu



obrázek IV.3-9: Autokovarianční funkce residuí

Pokusme se prověřit charakter residuí a posoudit do jaké míry splňují naše předpoklady a mohou ovlivňovat kvalitu nalezeného modelu. K tomu nám slouží další standardní nástroje residuální analýzy (obrázek IV.3-7 - obrázek IV.3-9). Časový vývoj residuí (obrázek IV.3-7) nám slouží k posouzení jejich charakteru. Na tomto grafu není patrný nějaký podstatný trend. Na dalším grafu (obrázek IV.3-8) jsou zobrazena residua v závislosti na řídicím signálu. Poznamenejme, že vstupní signál měl skokový charakter. Zde je patrná určitá závislost residuí na budícím signálu. Residua vykazují větší rozptyl a střední hodnotu pro nižší úroveň budícího signálu. To je patrně způsobeno nelinearitou tření v ložiskách. A na závěr vyhodnotíme autokovarianční funkci residuí (obrázek IV.3-9). Ta má příznivý průběh, i když poukazuje na určité nedostatky modelu. Na závěr lze říci, že tento model prošel všechny naše testy a je tedy vhodný pro použití v oblasti automatického řízení.

Jako další model jsme vyzkoušeli strukturu získanou z matematicko-fyzikální analýzy. Ačkoliv integrační či derivační metodu nelze na tento model použít, inicializační parametry pro iterační metodu optimalizace jsme získali na linearizované struktuře (bez Coulombova tření) integrační metodou. Model jsme uvažovali ve tvaru (IV-30):

$$y^{IV} + a_3 y''' + a_2 y'' + a_1 y' + a_0 y + D\eta = b_0 u \quad (\text{IV-30})$$

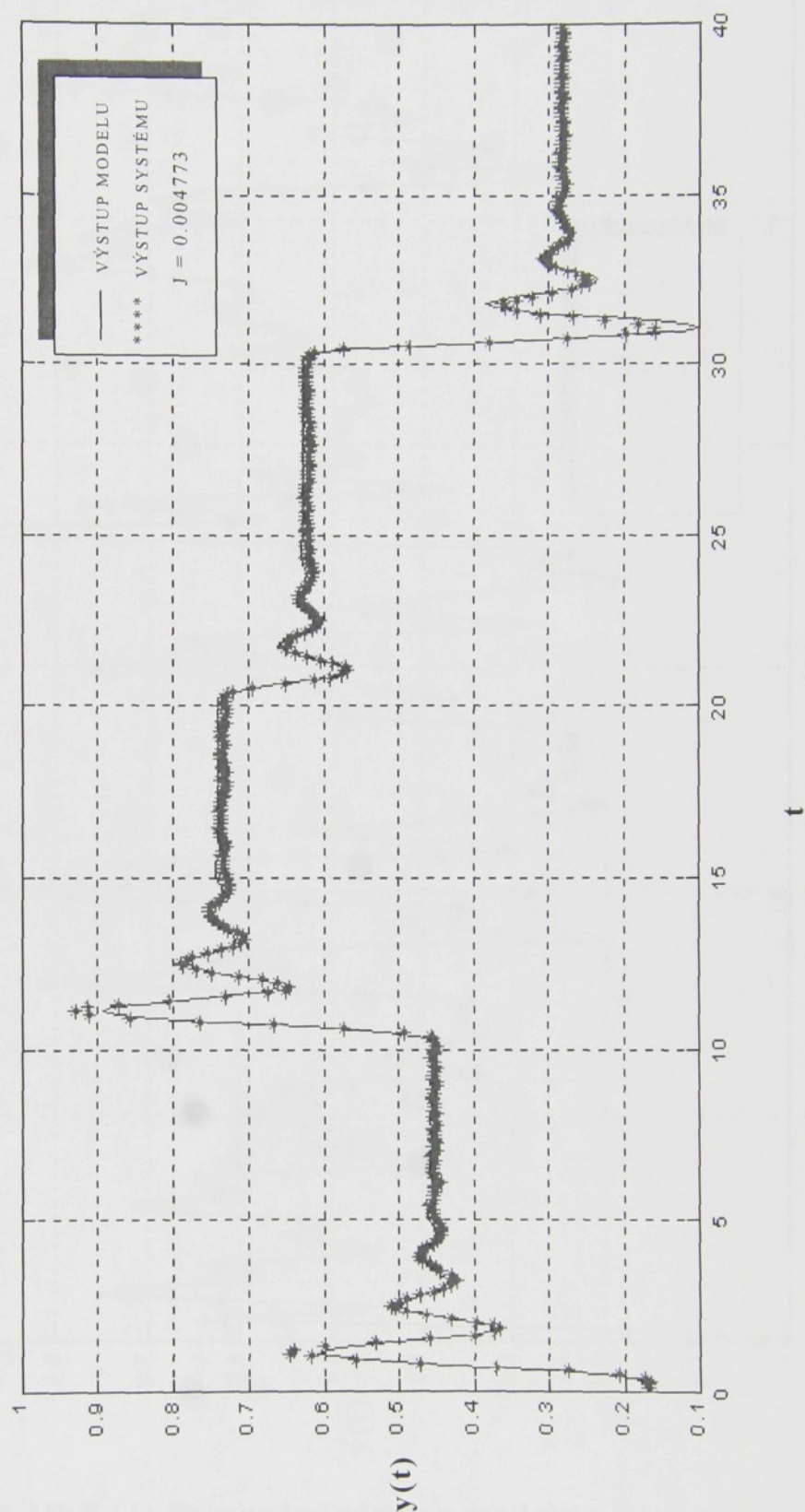
a nalezené parametry v rovnici (IV-30) jsou:

$$a_3 = 8.8748, a_2 = 68.8013, a_1 = 217.0087, a_0 = 739.0247$$

$$b_0 = 421.2441, D = 84.2488.$$

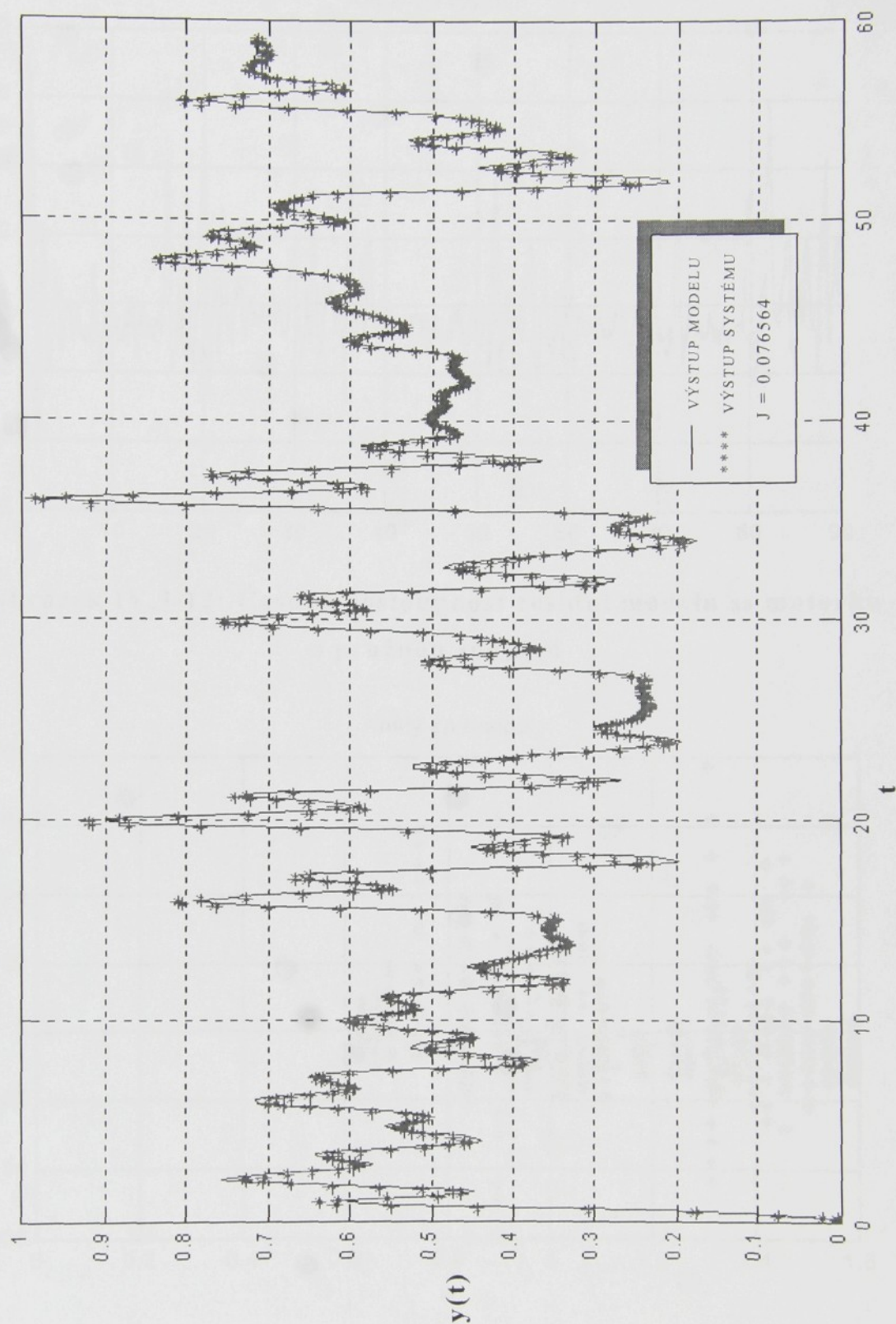
Jak jsme si už zvykli, provedme standardní analýzu kvality modelu. Podívejme se na porovnání výstupu nelineárního modelu a vyšetřovaného systému stejnosměrného motorku s pružnou spojkou (obrázek IV.3-10). Pro toto srovnání byla použita stejná data jako pro optimalizaci. Výstupy modelu a systému jsou velice podobné a i účelová funkce má vzhledem k průměrné hodnotě signálu a délce měření malou hodnotu. Na dalším obrázku (obrázek IV.3-11) je zobrazeno srovnání reakcí modelu a systému výměníku na jiný signál než byl použit pro optimalizaci. I toto srovnání dopadlo velmi dobře. Nalezený model pro použité rozsahy vstupních a výstupních veličin dobře reprodukuje chování studovaného systému.

Výhodnocení modelu, __Systém,*** Model s Coulom. třením

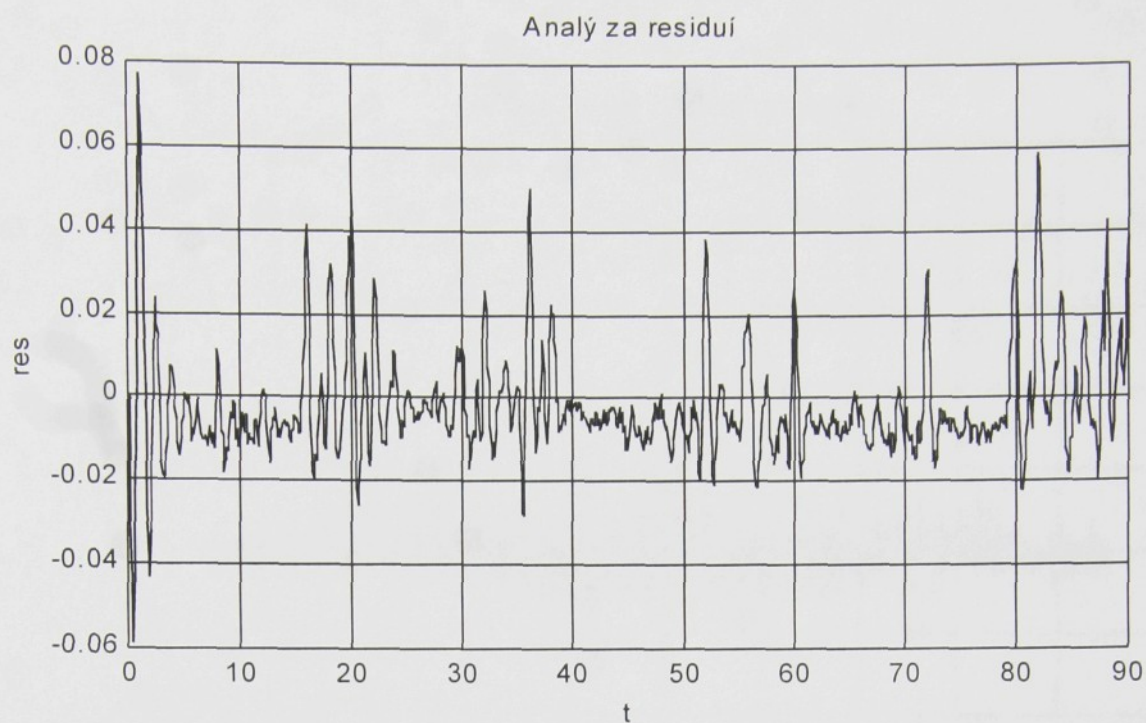


obrázek IV.3-10: Porovnání výstupu modelu a systému ss
motorku s pružnou spojkou na datech použitých pro optimalizaci

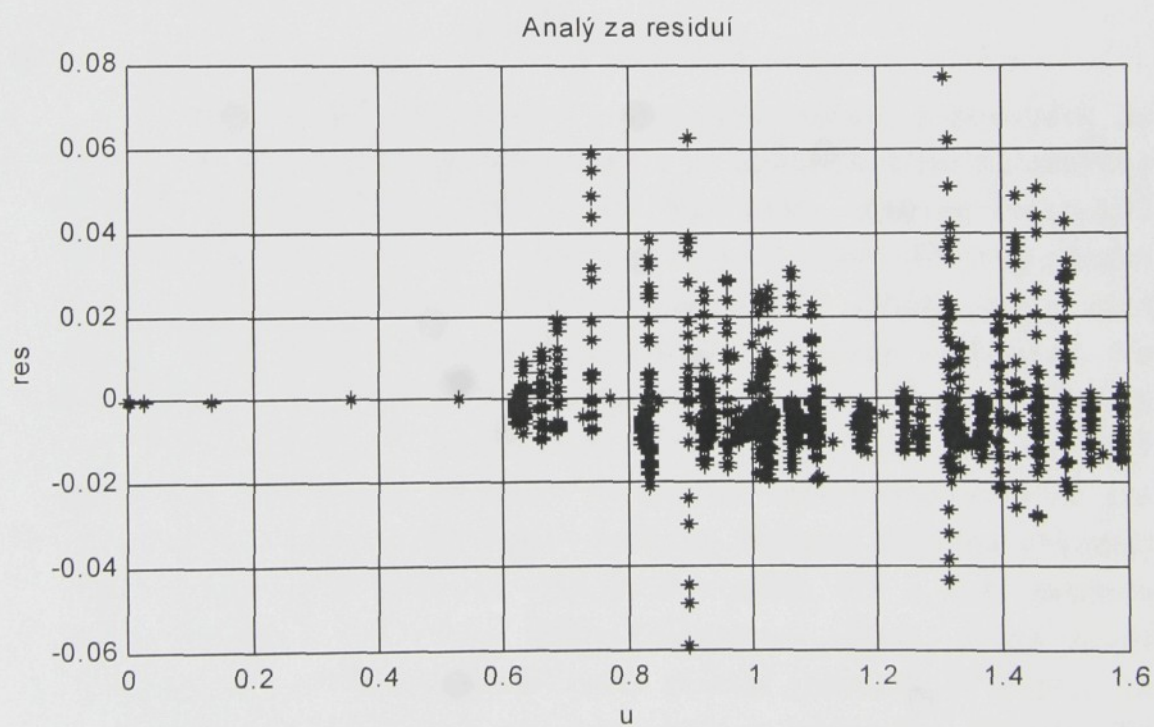
Vyhodnocení modelu, __Systém,*** Model s Coulom.



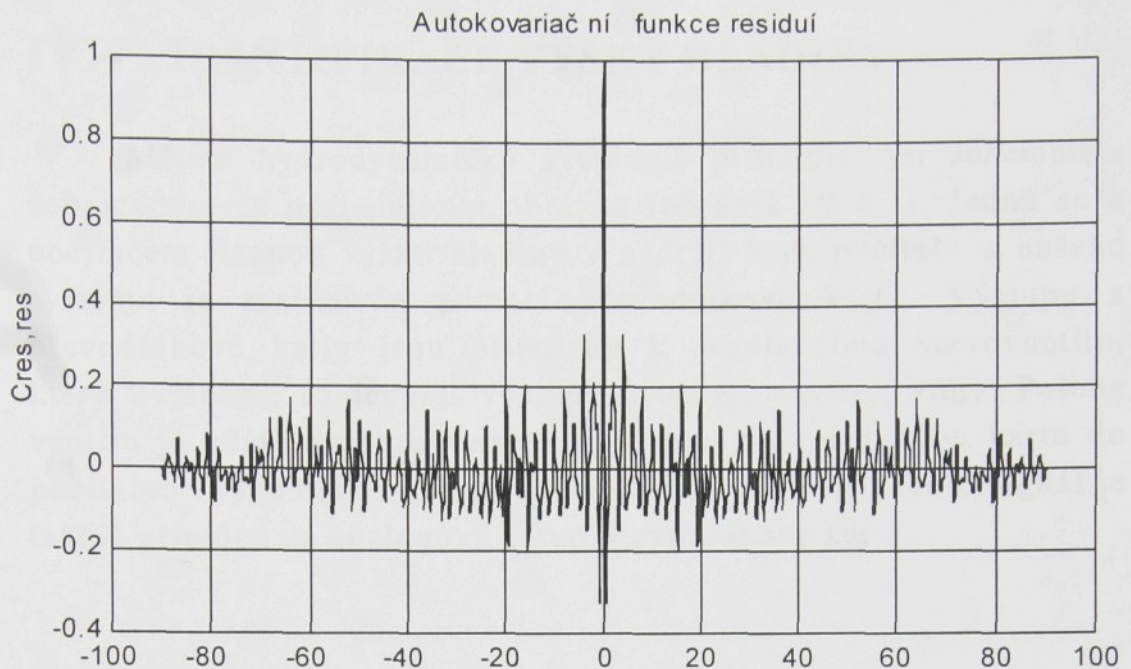
obrázek IV.3-11: Porovnání výstupu modelu a systému ss motorku s pružnou spojkou na jiných datech než byla použita pro optimalizaci



**obrázek IV.3-12: Časová posloupnost residuí modelu ss motorku
s pružnou spojkou**



obrázek IV.3-13: Graf residuí v závislosti na budícím signálu

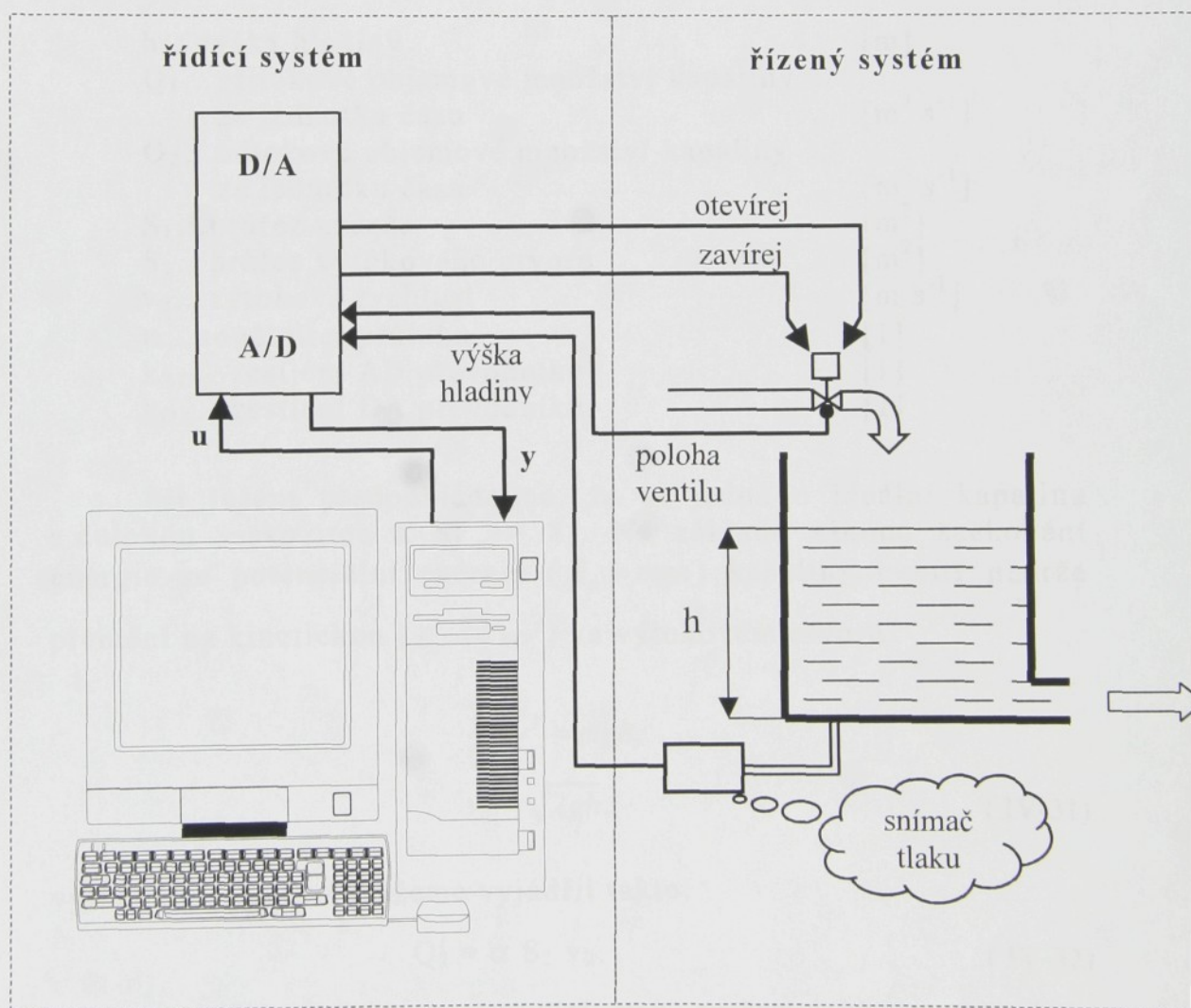


obrázek IV.3-14: Autokovarianční funkce residuí ss motorku s pružnou spojkou

Pokusme se nyní zhodnotit charakter residuí a posoudit, do jaké míry splňují naše předpoklady a ovlivňují kvalitu nalezeného modelu. K tomu nám slouží další standardní nástroje residuální analýzy (obrázek IV.3-12 - obrázek IV.3-14). Časový vývoj residuí nám slouží k posouzení jejich charakteru. Z tohoto grafu není zřejmý žádný trend, jen na počátku residua vykazují větší skok. Na dalším grafu (obrázek IV.3-13) jsou zobrazena residua v závislosti na řídicím signálu. Poznamenejme, že vstupní signál měl skokový charakter. Na tomto obrázku je patrná určitá struktura a lze usuzovat na závislost residuí na budícím signálu. Residua vykazují větší rozptyl okolo úrovně vstupního signálu 0,9 a 1,4, zatímco okolo úrovně 1,2 je tento rozptyl podstatně menší. A na závěr vyhodnoťme autokovarianční funkci residuí (obrázek IV.3-14). Ta má poněkud horší průběh než tomu bylo u lineárního modelu. To bude nejpravděpodobněji způsobeno odlišnou strukturou modelů. I tak lze ale prohlásit, že tento model prošel všechny naše testy a je tedy vhodný pro použití v oblasti automatického řízení.

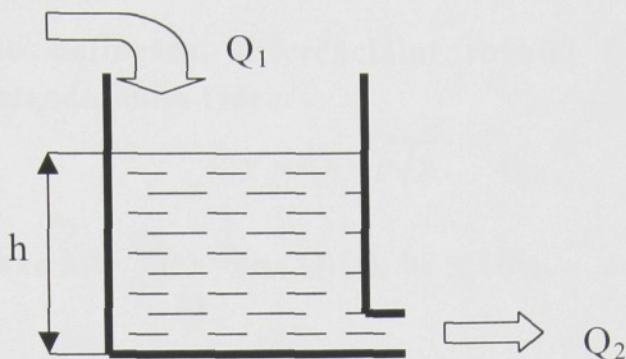
IV.4 IDENTIFIKACE VÝŠKY HLADINY

Mějme hydrodynamický systém s principiálním schématem zobrazeném na následujícím obrázku (obrázek IV.4-1). Jedná se o počítačem řízenou výšku hladiny v nádrži. Styk počítače a našeho systému je realizován pomocí převodníkové karty. Výstupy z převodníkové karty jsou připojeny k regulačnímu servoventilu, který ovládáme logickými výstupy (otvírej, zavírej, stůj). Poloha ventilu je odměřována a vedena zpět přes převodníkovou kartu do počítače. Výška hladiny je snímána snímačem tlaku a tento signál je taktéž připojen na analogový vstup převodníkové karty.



obrázek IV.4-1: Výška hladiny - schéma systému

Pokusme se pomocí matematicko-fyzikální analýzy výše popsaného systému odvodit diferenciální rovnice vyjadřující dynamickou závislost výšky hladiny h kapaliny ve válcové nádobě s volným odtokem na přitékajícím množství kapaliny Q_1 (viz.obrázek IV.4-2).



obrázek IV.4-2: Principiální schéma modelu výšky hladiny

Definujme:

h ...výška hladiny	[m]
Q_1 ...přítokové objemové množství kapaliny za jednotku času	$[m^3 s^{-1}]$
Q_2 ...odtokové objemové množství kapaliny za jednotku času	$[m^3 s^{-1}]$
S_1 ...průřez nádrže	$[m^2]$
S_2 ...průřez výtokového otvoru	$[m^2]$
v_2 ...výtoková rychlost	$[m s^{-1}]$
α ...součinitel průtoku	[1]
k_{AD} ...zesílení AD převodníku	[1]
k_{DA} ...zesílení DA převodníku	[1]

Při řešení předpokládejme, že se jedná o ideální kapalinu s nulovou viskozitou a $S_1 \gg S_2$. Na základě zákona zachování energie se potenciální energie ($E_p = mgh$) kapaliny uvnitř nádrže přemění na kinetickou ($E_k = \frac{1}{2}mv^2$) ve výtokovém otvoru:

$$\begin{aligned} \frac{1}{2}mv^2 &= mgh, \\ v_2 &= \sqrt{2gh}, \end{aligned} \quad (IV-31)$$

výtokové množství můžeme vyjádřit takto:

$$Q_2 = \alpha S_2 v_2. \quad (IV-32)$$

A jestliže napíšeme rovnici pro změnu objemu kapaliny v nádrži, pak platí:

$$\frac{dV}{dt} = S_1 \frac{dh}{dt} = Q_1 - Q_2 = Q_1 - \alpha S_2 \sqrt{2g} \sqrt{h}, \quad (\text{IV-33})$$

$$h' + \frac{\alpha S_2 \sqrt{2g}}{S_1} \sqrt{h} = \frac{1}{S_1} Q_1. \quad (\text{IV-34})$$

A tuto nelineární diferenciální rovnici (IV-34) můžeme vyjádřit ve standardním tvaru:

$$y' = b_0 u - p \sqrt{y}, \quad (\text{IV-35})$$

$$\text{kde: } y(t) = k_{AD} h(t), \quad u(t) = k_{DA} Q_1(t), \quad b_0 = 1/S_1, \quad p = \frac{\alpha S_2 \sqrt{2g}}{S_1}.$$

Z hlediska pomalé dynamiky výšky hladiny lze zanedbat dynamiku akčního členu (ventil) a snímačů. Model byl tedy uvažován ve tvaru totožném s (IV-35).

Nalezené parametry integrační metodou jsou:

$$b_0 = 0.147012, \quad p = 0.012614.$$

Nalezené parametry derivační metodou jsou:

$$b_0 = 0.145316, \quad p = 0.012087.$$

Parametry nalezené těmito metodami jsou velice podobné a i odezvy těchto modelů jsou téměř identické, stejně je tomu pak i s následným vyhodnocením kvality modelu, proto jsou zde uvedeny pouze výsledky srovnání modelu a systému pro model nalezený integrační metodou.

Toto porovnání výstupu modelu získaného integrační metodou a vyšetřovaného systému výšky hladiny kapaliny v nádrži je uvedeno na následujícím grafu (obrázek IV.4-3). Pro toto srovnání byla použita stejná data jako pro optimalizaci. Výstupy modelu a systému jsou velice podobné v celé šířce výstupního signálu y a i účelová funkce má vzhledem k průměrné hodnotě signálu a délce měření malou hodnotu. Na dalším obrázku (obrázek IV.4-4) je zobrazeno srovnání reakcí modelu a systému výměníku na jiný signál než byl použit pro optimalizaci. I toto srovnání dopadlo velmi dobře a není patrný znatelný posun k horšímu, i když, jak se dalo předpokládat, kvadratická účelová funkce má o něco vyšší hodnotu. Nalezený model pro použité rozsahy vstupních a výstupních veličin dobře reprodukuje chování studovaného systému.

Nakonec ještě vyzkoušejme nalézt parametry modelu stejné struktury viz.(IV-35) iterační metodou, která byla inicializována parametry získanými pomocí derivační metody.

Nalezené parametry iterační metodou:

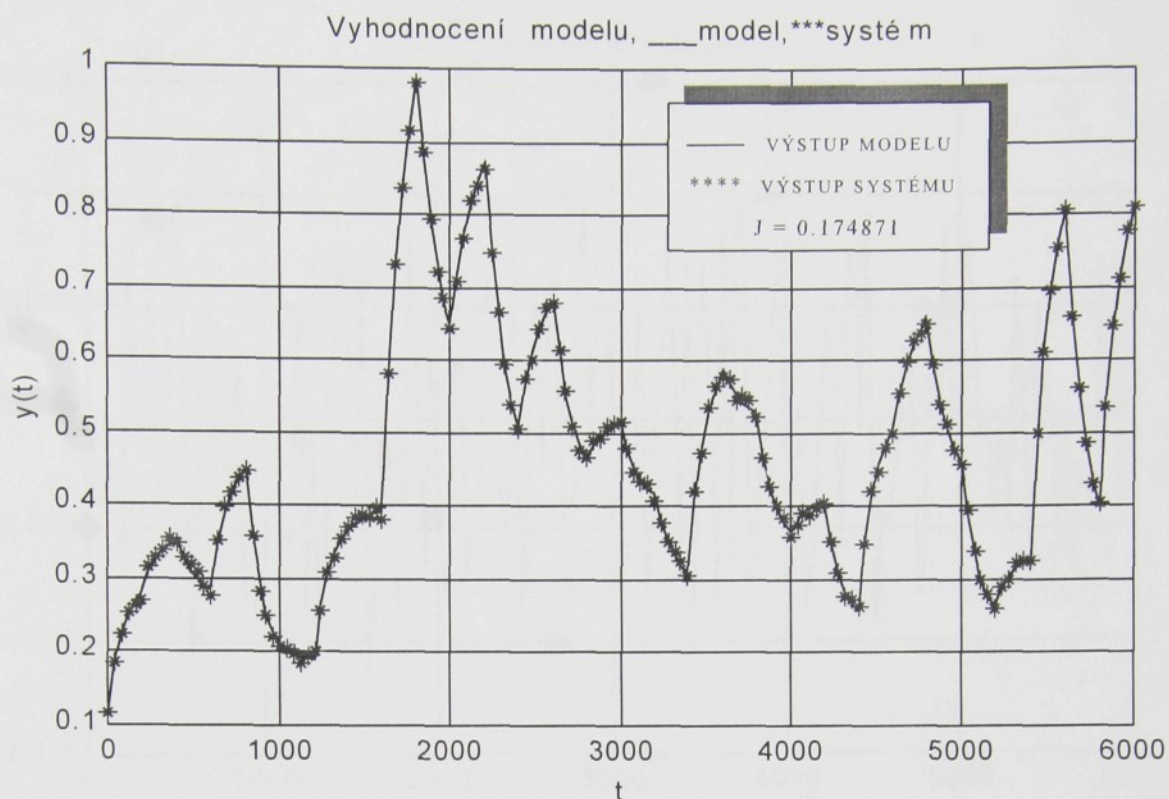
$$b_0 = 0.147293, p = 0.012114.$$

Porovnání výstupu modelu získaného iterační metodou a vyšetřovaného systému výšky hladiny kapaliny v nádrži je uvedeno na následujícím grafu (obrázek IV.4-8). Pro toto srovnání byla použita stejná data jako pro optimalizaci. Na dalším obrázku (obrázek IV.4-9) je zobrazeno srovnání reakcí modelu a systému výměníku na jiný signál než byl použit pro optimalizaci. Tato srovnání dopadla velice podobně jako o modelu nalezeného integrační metodou, takže i pro tento model platí vše, co bylo řečeno v předchozím odstavci.

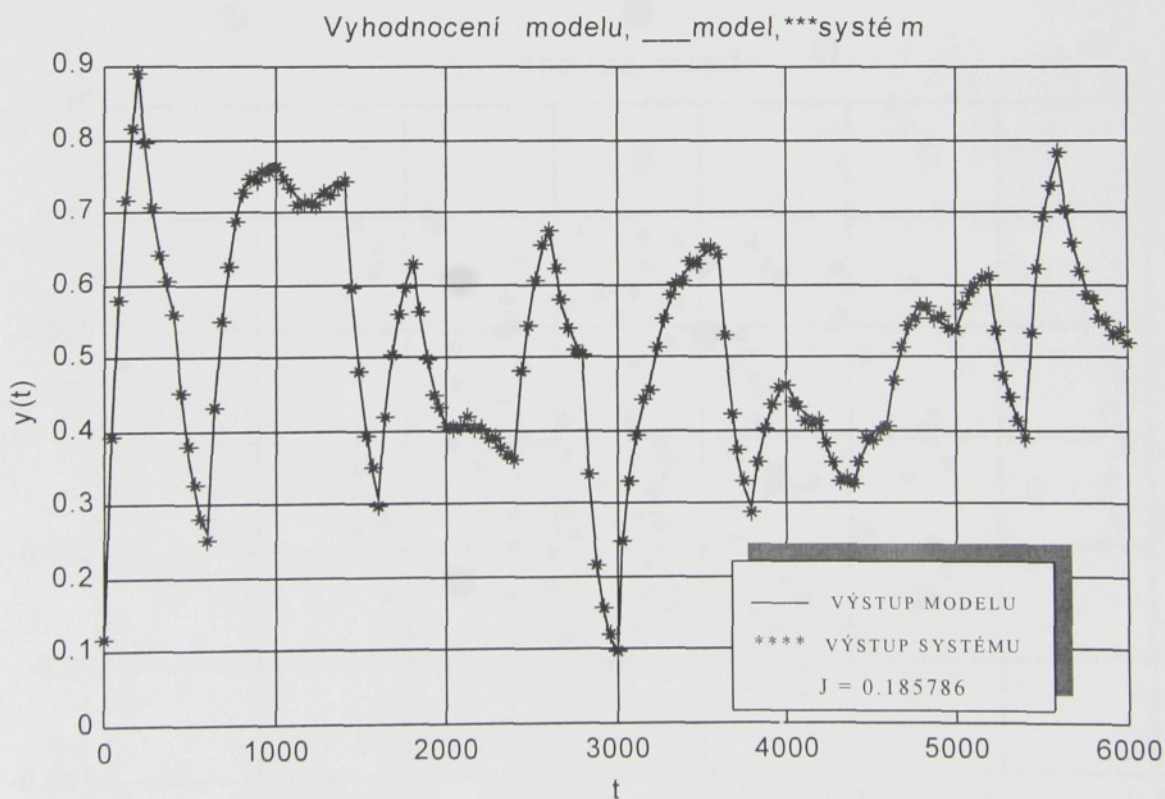
Porovnejme hodnoty kvadratické účelové funkce pro jednotlivé metody a tak zhodnoťme jak si vedly (tabulka IV.4-1). Z této tabulky vidíme, že nejlépe si vedla iterační metoda a nejhůře integrační, i když jsou mezi nimi jen nepatrné rozdíly.

<i>Metoda</i>	<i>Hodnota účelové funkce J(p)</i>
<i>Integrační metoda</i>	0.184958
<i>Derivační metoda</i>	0.174872
<i>Iterační metoda</i>	0.171594

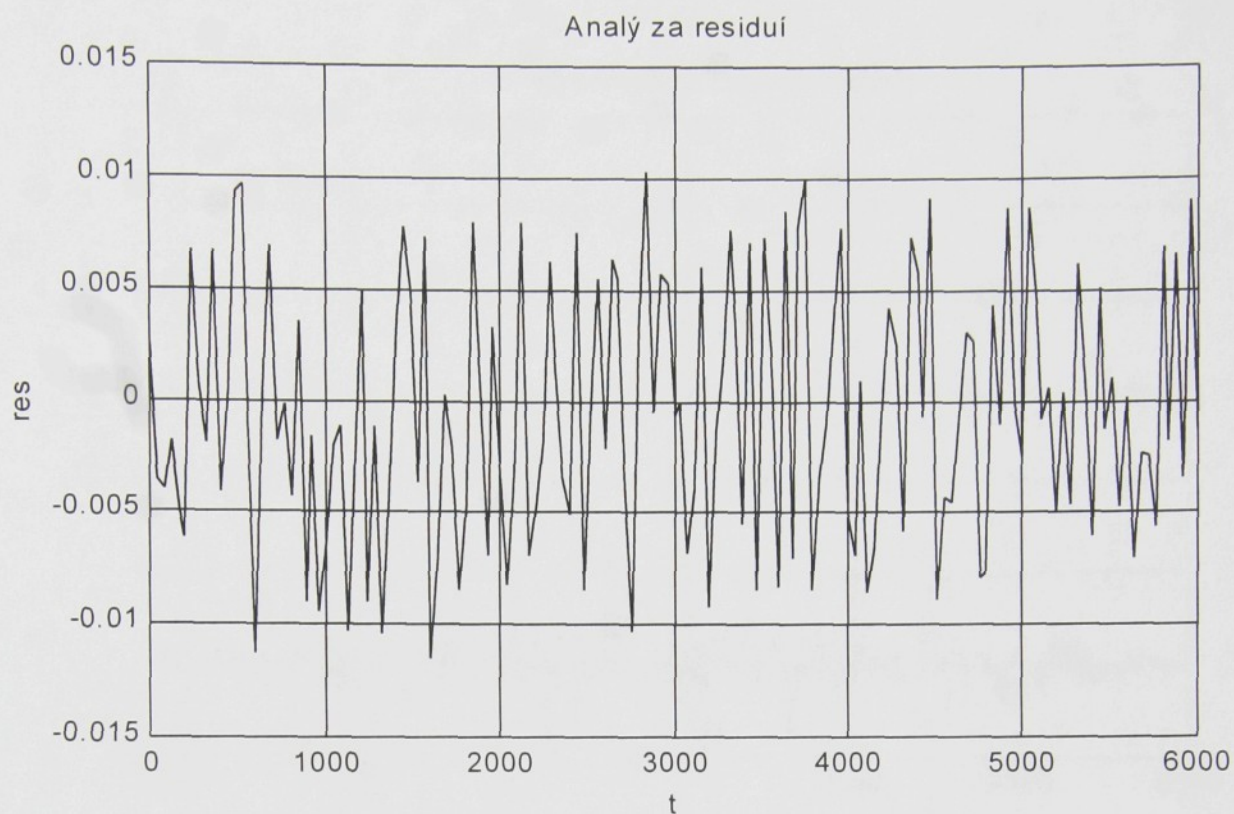
tabulka IV.4-1: Hodnoty účelové funkce pro systém výšky hladiny



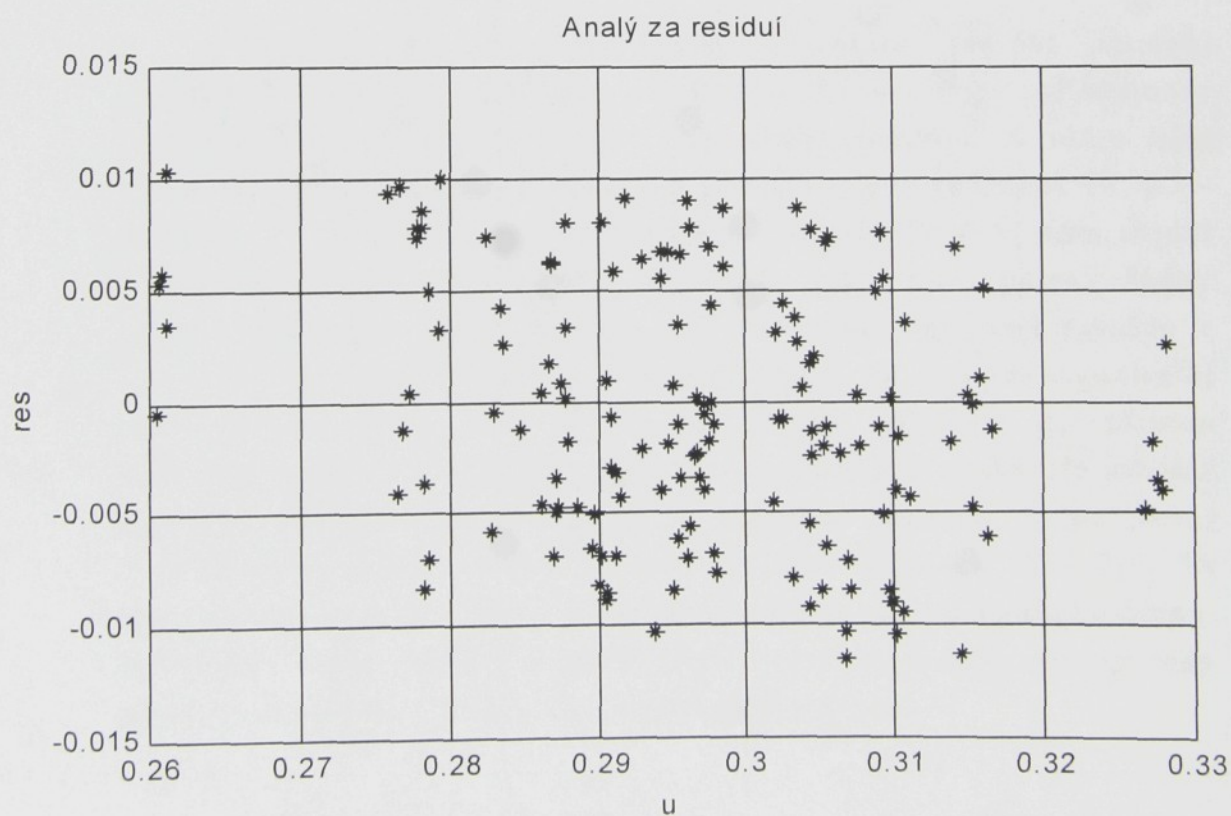
obrázek IV.4-3: Porovnání výstupu modelu a systému výšky hladiny na datech použitých pro optimalizaci



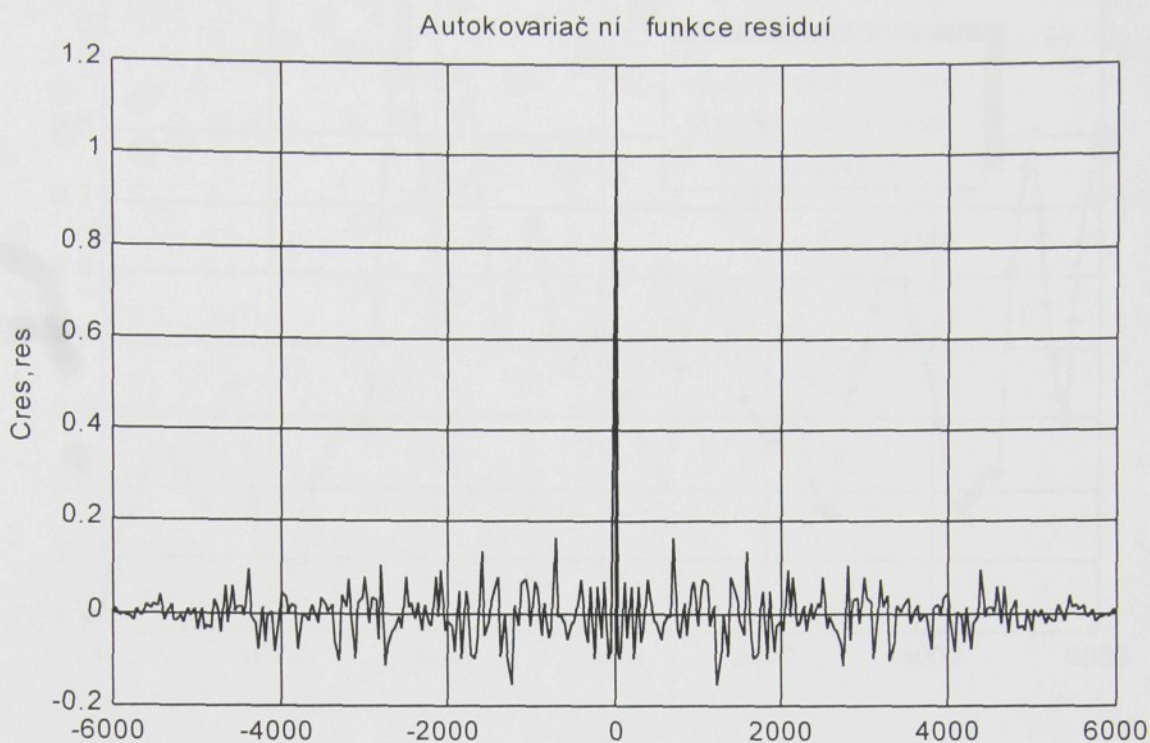
obrázek IV.4-4: Porovnání výstupu modelu a systému výšky hladiny na jiných datech než byla použita pro optimalizaci



obrázek IV.4-5: Časová posloupnost residuí modelu výšky hladiny

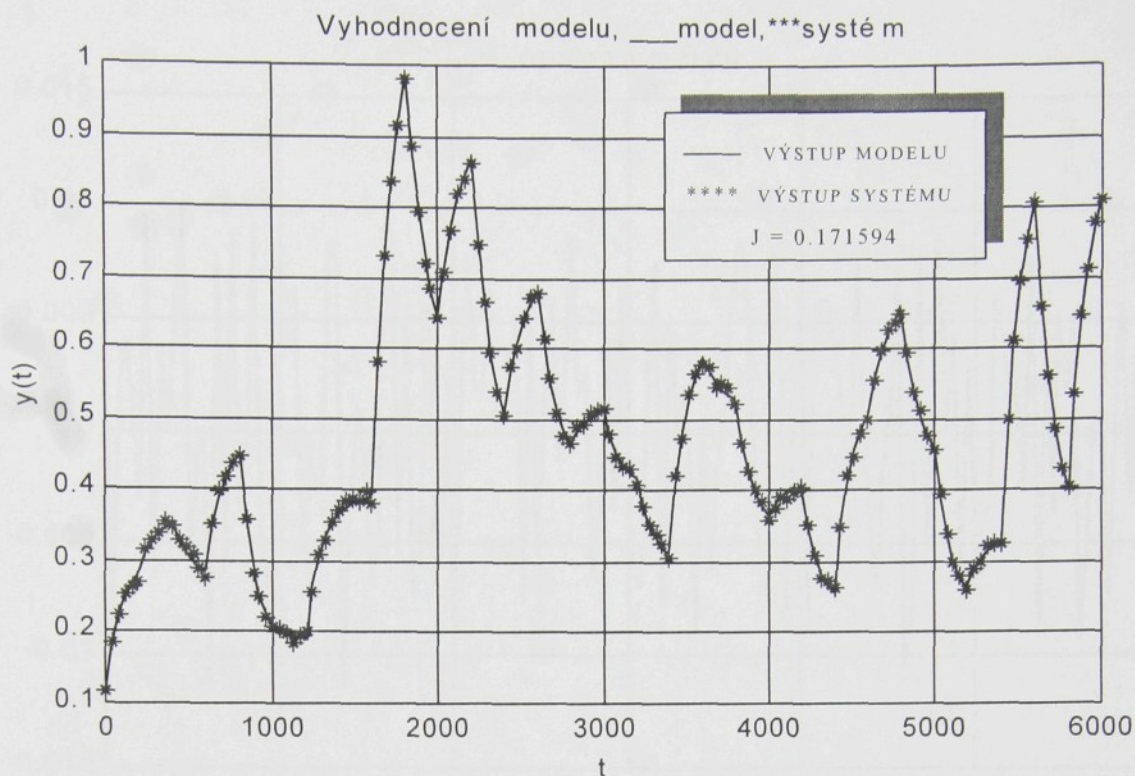


obrázek IV.4-6: Graf residuí v závislosti na budícím signálu

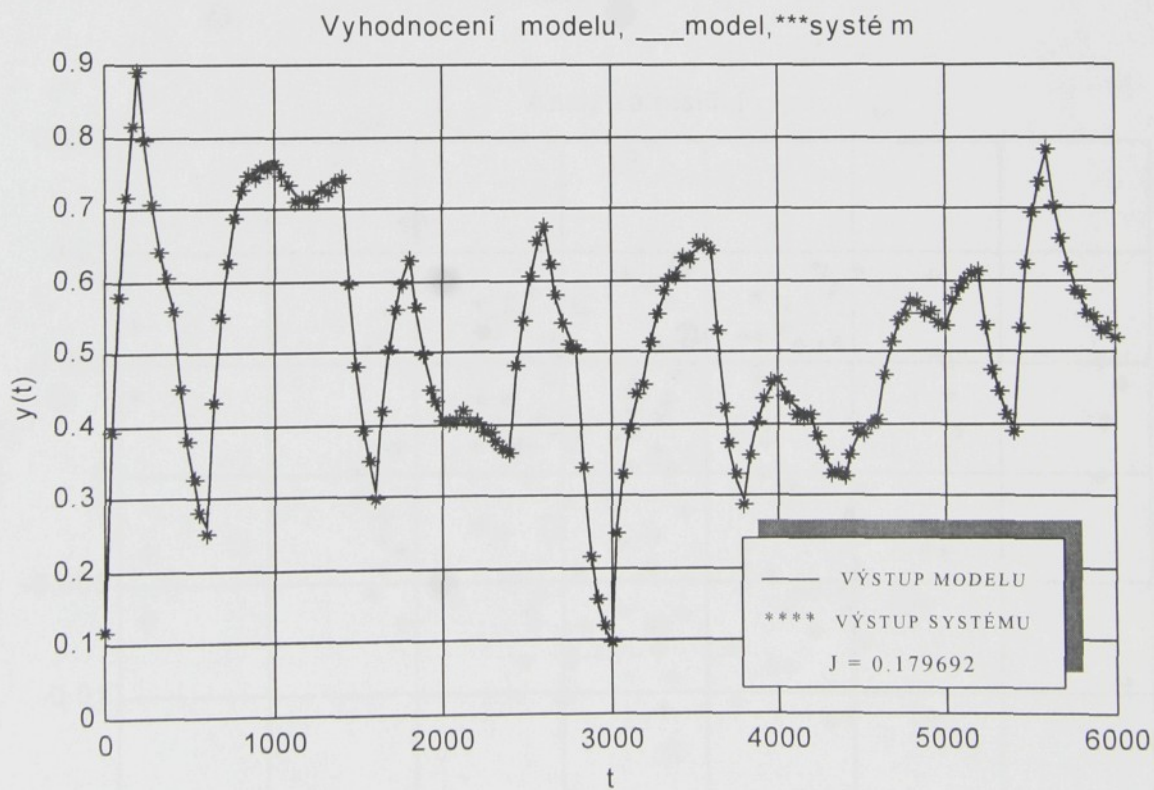


obrázek IV.4-7: Autokovarianční funkce residuí systému výšky hladiny

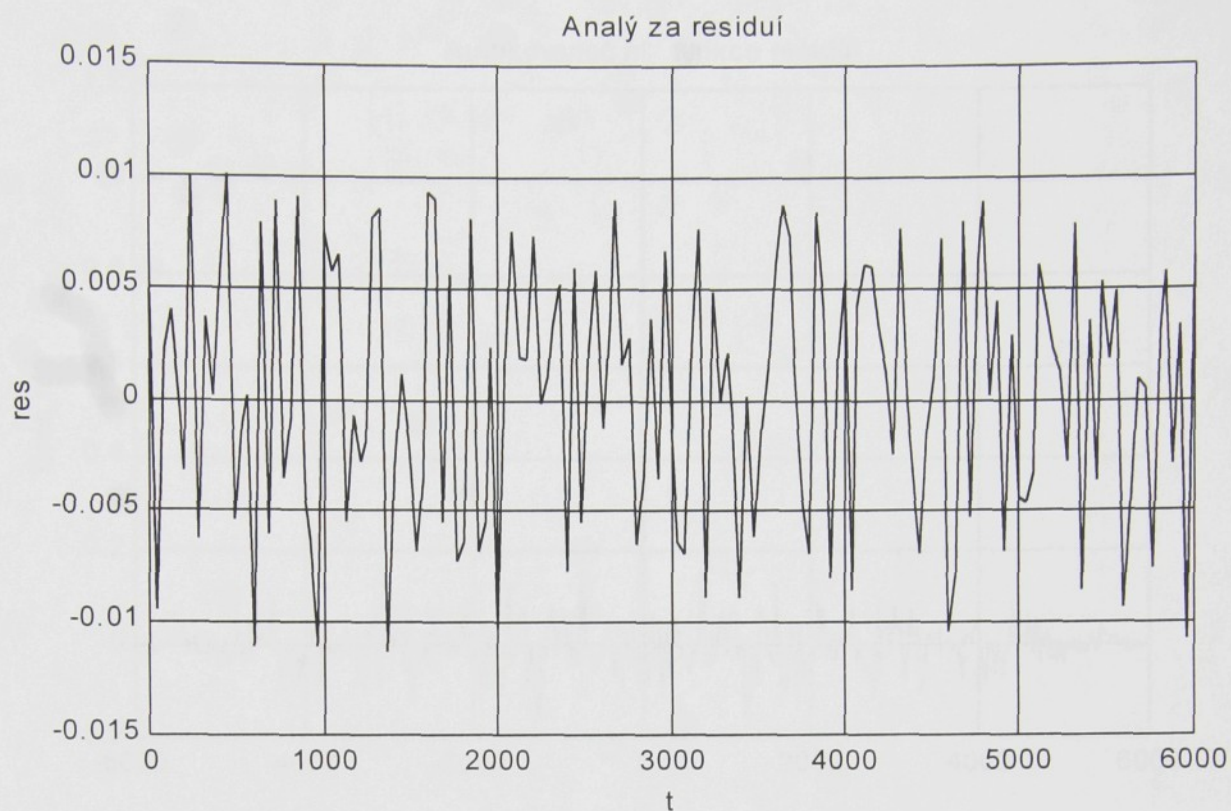
Pokusme se nyní zhodnotit charakter residuí modelu nalezeného integrační metodou a posoudit, do jaké míry splňují naše předpoklady a ovlivňují kvalitu nalezeného modelu. K tomu nám slouží další standardní nástroje residuální analýzy (obrázek IV.4-5 - obrázek IV.4-7). Časový vývoj residuí (obrázek IV.4-5) nám slouží k posouzení jejich charakteru. Z tohoto grafu není patrný žádný trend. Na dalším grafu (obrázek IV.4-6) jsou zobrazena residua v závislosti na řídicím signálu. Poznamenejme, že vstupní signál měl skokový charakter. Ani zde naštěstí nelze usuzovat na přímou závislost residuí na budícím signálu; jen se lehce snižovala střední hodnota residuí při vyšší hodnotě budícího signálu. A na závěr vyhodnotíme autokovarianční funkci residuí (obrázek IV.4-7). Ta má příznivý průběh blížící se autokov. funkci nekorelovaného šumu. Na závěr lze říci, že tento model prošel všechny naše testy a je tedy vhodný pro použití v oblasti automatického řízení.



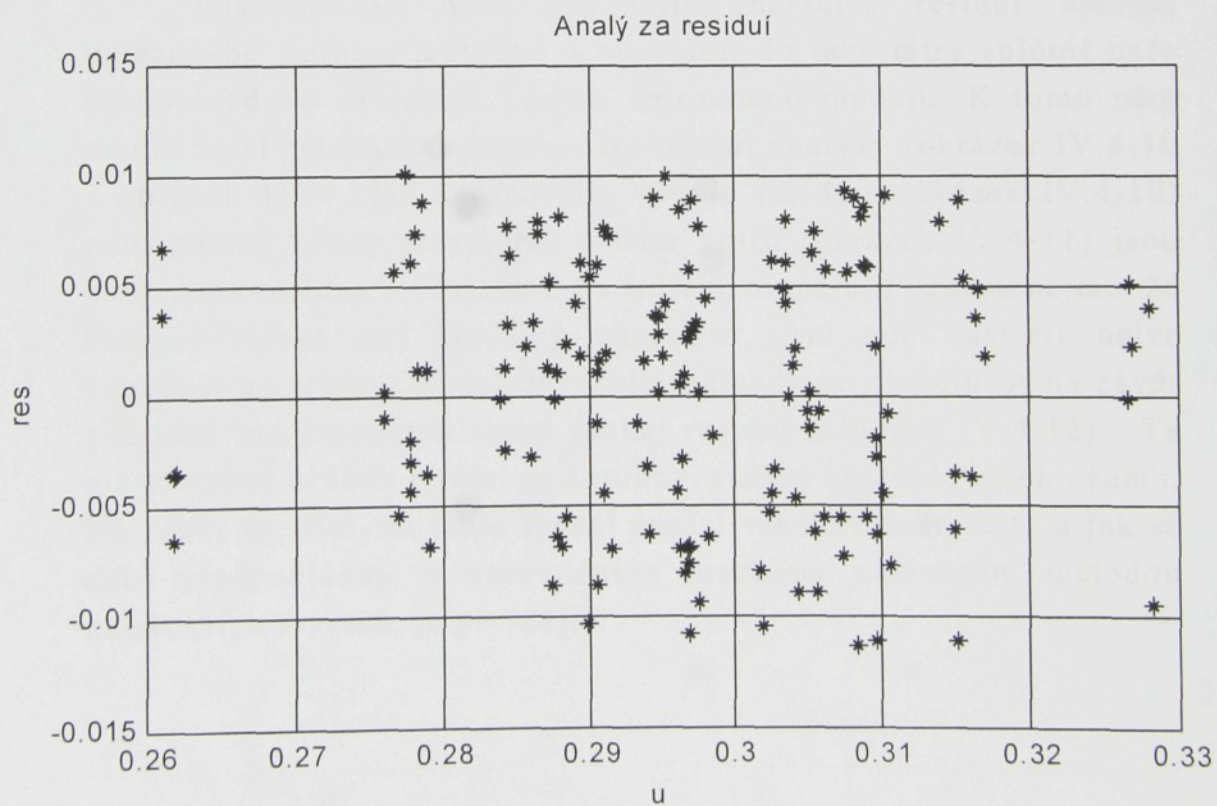
obrázek IV.4-8: Porovnání výstupu modelu a systému výšky hladiny na datech použitých pro optimalizaci



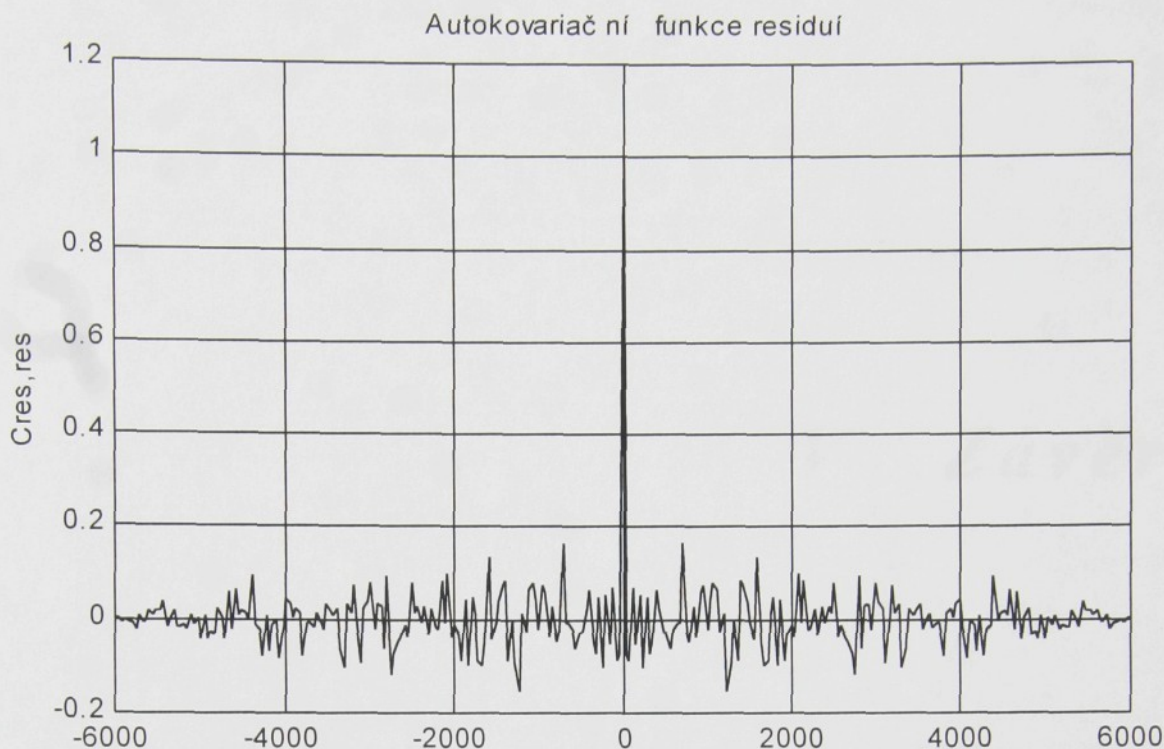
obrázek IV.4-9: Porovnání výstupu modelu a systému výšky hladiny na jiných datech než byla použita pro optimalizaci



obrázek IV.4-10: Časová posloupnost residuí modelu výšky hladiny



obrázek IV.4-11: Graf residuí v závislosti na budícím signálu



obrázek IV.4-12: Autokovarianční funkce residuí systému výšky hladiny

Pokusme se nyní zhodnotit charakter residuí modelu nalezeného iterační metodou a posoudit, do jaké míry splňují naše předpoklady a ovlivňují kvalitu nalezeného modelu. K tomu nám slouží další standardní nástroje residuální analýzy (obrázek IV.4-10 - obrázek IV.4-12). Z časového vývoje residuí (obrázek IV.4-10) není patrný žádný trend. Na dalším grafu (obrázek IV.4-11) jsou zobrazena residua v závislosti na řídicím signálu. Poznamenejme, že vstupní signál měl skokový charakter. Ani zde naštěstí nelze usuzovat na přímou závislost residuí na budícím signálu. A na závěr vyhodnoťme autokovarianční funkci residuí (obrázek IV.4-12). Ta má příznivý průběh blížící se autokov. funkci nekorelovaného šumu. Na závěr lze říci, že tento model prošel všechny naše testy a jak se dalo předpokládat, v porovnání s modelem nalezeným metodou nejmenších čtverců, je přesnější.

V Závěr

Úkolem této práce bylo nalézt a vyzkoušet postupy identifikace deterministických lineárních a nelineárních parametrizovaných modelů ve spojité časové oblasti. Je zde uvedeno odvození dvou metod (integrační a derivační) pracujících s modely lineárními v parametrech. Pro tyto dvě metody lze použít metodu nejmenších čtverců, což značně zrychluje celý výpočet. Nevýhodou těchto metod je, že při řešení nahrazujeme vnitřní proměnné modelu, který je nelineární v parametrech, korespondujícími hodnotami ze změřených dat, což může v některých případech (např. pro špatně zvolený model, či systém ovlivněný poruchou) vést k nesmyslnému výsledku. Pro tyto dvě metody je odvozeno rozšíření na soustavy s nenulovými počátečními podmínkami. Zatímco u integrační metody nenulové počáteční podmínky zvyšují počet hledaných parametrů v závislosti na řádu modelu, hledání parametrů derivační metodou tímto není nijak ovlivněno. Tyto dvě metody byly podrobeny analýze citlivosti kvality nalezeného modelu v závislosti na působící poruše a řádu modelu. Zatímco šum o vyšších frekvencích nemá tak výrazný vliv na nalezený model, šum o nízkých frekvencích již při menších úrovních působí značně destruktivně na výsledný model. Částečným řešením je rozšíření modelu o parametry trendu. Negativní vliv šumu je dále umocňován řádem identifikované soustavy (respektive modelu). Obecně si lépe vedla integrační metoda v porovnání s derivační a to platí i pro

velikost vzorkovací periody, protože derivační metoda pro stejnou přesnost potřebuje menší vzorkovací periodu.

Další metodou, která se dostala do středu našeho zájmu, je metoda kvazilinearizace. Tato metoda již patří do třídy iteračních metod a pomocí náhrady nelineární funkce dvěmi prvními členy Taylorova rozvoje je schopna hledat parametry nelineárních modelů. Značným omezením této metody je nutná znalost prvních parciálních derivací dle vektoru hledaných parametrů. S výhodou byla vyzkoušena kombinace integrační metody a kvazilinearizace. Toto spojení sníží počet iterovaných parametrů pro většinu případů hledání modelů pro použití v oblasti automatického řízení.

Poslední prověřovaný postup završující skupinu metod ve spojitě časové oblasti nám umožňuje hledat parametry volně parametrizovatelných struktur dle volně definovaného kritéria. Tento postup využívá univerzálních iteračních metod a byl úspěšně implementován v softwarovém prostředí MATLAB-SIMULINK. Značnou úsporu času výpočtu lze ve většině případů dosáhnout kombinací této metody a integrační (respektive derivační) metody, kdy integrační metoda je použita na hledání parametrů lineárního subsystému, ostatní parametry jsou počítány iteračním algoritmem.

V závěru této disertační práce byla demonstrována použitelnost těchto metod na několika skutečných systémech. Nalezené modely byly následně podrobeny kvalitativní analýze a jednotlivé metody a přístupy byly porovnány mezi sebou. Z výsledků je zřejmé, že metody identifikace ve spojitě časové oblasti mají značný potenciál pro nalezení kvalitních modelů především tam, kde něco víme o podstatě identifikovaného procesu a pouze hledáme neznámé parametry. Tyto metody si ale určitě nekladou za cíl konkurovat metodám v diskrétní oblasti (v důsledku jejich odlišnosti to ani není možné), ale spíše je mají doplňovat a v některých případech i třeba alternovat. Dalším doporučením vyplývajícím z této práce je doplňování metod používajících metodu nejmenších čtverců (ty lze použít pro počáteční odhad parametrů iterační metody) některou z iteračních metod. Tímto modifikovaným přístupem se celý postup identifikace stává daleko robustnějším vůči různým vnějším negativním vlivům.

SEZNAM LITERATURY

- [BEL] Bellman, Richard Ernest, Roth, R.: Quasilinearization and the identification problem, Singapore, World Scientific, 1983.
- [BOS] Bosch, P. P. J., Klauw, A. C.: Modeling, identification, and simulation of dynamical systems, CRC Press, 1994.
- [FON] Forsman, K., Ljung, L.: On the Deadzone in System Identification, technical report, Dept. of EE. Linköping University, 1991.
- [FOR] Forssell, U., Ljung, L.: Identification for Control: Some Results on Optimal Experiment Design, technical report, Dept. of EE. Linköping University, 1998.
- [FOR1] Forssell, U.: Identification for control: unification of some existing schemes, technical report, Dept. of EE. Linköping University, 1997.
- [FOR2] Forssell, U, Lindskog, P.: Combining Semi-physical and Neural Network Modeling: An Example of Its Usfulness, Proceedings of the 11th IFAC Symposium on System Identification, pp 795-798, 1997.
- [GUN] Gunnarsson, S.: Combining Tracking and Regularization in Recursive Least Squares Identification, technical report, Dept. of EE. Linköping University, 1996.
- [GUN1] Gunnarsson, S., Hjalmarsson, H.: Some Aspects of Iterative Identification and Control Design Schemes, technical report, Dept. of EE. Linköping University, 1994.
- [HAN1] Hanuš Bořivoj - Alexík, M. – Bíla, J. – Kroc, L. – Krupka, Z. : Teorie automatického řízení II, Liberec, VŠST Liberec, 1985.
- [HAN2] Hanuš Bořivoj – Balátě, J. – Švarc, I. – Zikeš, F.: Teorie automatického řízení I, 1. a 2. část, Liberec, VŠST Liberec, 1982.
- [HAN3] Hanuš Bořivoj – Balda, M.: Základy technické kybernetiky I, Liberec, VŠST Liberec, 1989.
- [HAN4] Hanuš Bořivoj – Balda, M.: Základy technické kybernetiky II. část, Brno, VUT Brno, 1981.
- [HEI] Heij, C.: Deterministic identification of dynamical systems, New York, Springer-Verlag, 1989.
- [HOR] Horáček, P., Fuka, J.: Systémy a modely, ČVUT, Praha, 1996.
- [CHE] Ch'en, Han-fu, Guo, L.: Identification and stochastic adaptive control, Boston, Birkhäuser, 1991.

- [JAJ] Janeček Josef – Modrlák, O.: Základy technické kybernetiky, příklady, Liberec, VŠST Liberec 1990.
- [JOH] John, J., Horáček, P.: Identifikace a modelování, ČVUT, Praha, 1982.
- [LIN] Lindskog, P.: Methods, Algorithms and Tools for System Identification Based on Prior Knowledge, PhD. Thesis, Linköping University, 1996.
- [LIN1] Lindskog, P.: Fuzzy identification from a grey box modeling point of view, technical report, Dept. of EE. Linköping University, 1996.
- [LJU1] Ljung Lennart: System Identification Toolbox : For Use with MATLAB, Natic, Math Works, 1993.
- [LJU2] Ljung Lennart: System identification : Theory for the user, Englewood Cliffs, PTR Prentice Hall, 1987.
- [LJU3] Ljung Lennart: Identification for Control -- What is there to learn?, Learning, Control and Hybrid Systems, st. 207-221, 1998.
- [LJU4] Ljung Lennart: Developments for the System Identification Toolbox for MATLAB, technical report, Dept. of EE. Linköping University, 1999.
- [LJU5] Ljung Lennart,: Model Validation and Model Error Modeling , technical report, Dept. of EE. Linköping University, 1999.
- [LJU6] Ljung Lennart, Guo, L.: Classical Model Validation for Control Design Purposes, Mathematical Modelling of Systems, Vol 3, st. 27-42, 1997.
- [LJU7] Ljung Lennart, Guo, L.: The Role of Model Validation for Assessing the Size of the Unmodeled Dynamics, IEEE Trans. Automatic Control, Vol AC-42, st. 1230-1240, 1997.
- [LJU8] Ljung Lennart,: Development of system identification, technical report, Dept. of EE. Linköping University, 1996.
- [LJU9] Ljung Lennart,: System identification through the eyes of model validation, technical report, Dept. of EE. Linköping University, 1995.
- [LJU10] Ljung Lennart,: Some Results on Identifying Linear Systems Using Frequency Domain Data, Proc. of the 32th IEEE Conference on Decision and Control, st. 3534-3538, 1993.
- [LJU11] Ljung Lennart, Glad, T., Andersson, T.: Identifiability implies robust identifiability, technical report, Dept. of EE. Linköping University, 1993.
- [LJU12] Ljung Lennart, Sjöberg, J.: A system identification perspective on neural nets, technical report, Dept. of EE. Linköping University, 1992.

- [LJU13] Ljung Lennart, Sjöberg, J., McKelvey, T.: On the use of regularization in system identification, technical report, Dept. of EE. Linköping University, 1992.
- [MAN] Maňas Miroslav: Optimalizační metody, Praha, SNTL, 1979.
- [MAT] Matko, D., Zupančič, B., Karba, R.: Simulation and Modelling of Continuous Systems : A Case Study Approach, Englewood Cliffs: Prentice-Hall, 1992.
- [McK] McKelvey, T.: Identification of State-Space Models from Time and Frequency Data, PhD. Thesis, Linköping University, 1995.
- [MOD1] Modrlák, O.: Teorie automatického řízení, Liberec, 1992.
- [MOD2] Modrlák, O.: Dynamic Systems and Feedback Control : An Introduction, Liberec, 1992.
- [MOD3] Modrlák, O.: Theory of Control, Liberec, 1992.
- [NIU] Niu, S., Ljung, L., Björk, Å.: Decomposition Methods for Solving Least-Squares Parameter Estimation, IEEE Trans. Signal Processing, Vol 44, st. 2847, 1996.
- [NOR] Norton, J. P.: An introduction to identification, New York, Academic Press, 1985.
- [OLE] Olehla, M.: Identifikace technologických soustav : Studijní podklady pro obor ASŘ ve strojírenství, TU Liberec, Liberec, 1997.
- [PRE] Press William H.: Numerical recipes in C : The art of scientific computing, Cambridge, Cambridge University, 1992.
- [SAH] Saha, Dines Chandra: Identification of continuous dynamical systems : the Poisson Moment Functional (PMF) approach, Springer-Verlag, 1983.
- [SJO] Sjöberg, J.: Non-Linear System Identification with Neural Networks, PhD. Thesis, Linköping University, 1995.
- [SJO1] Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Deylon, B., Glorennec, P., Hjalmarsson, H., Juditsky, A.: Nonlinear black-Box Modeling in System Identification: a Unified Overview, technical report, Dept. of EE. Linköping University, 1995.
- [SJO2] Sjöberg, J.: Some Examples of Identification with Neural Networks, technical report, Dept. of EE. Linköping University, 1994.
- [SLO] Slotine, J., Weiping L.: Applied Nonlinear Control, Englewood Cliffs : Prentice-Hall, 1991.

[SOD] Söderström Torsten – Stoica, P.: System Identification, Hemel Hempstead: Prentice Hall, 1989.

[SOU] Soukup Jan: Identifikace soustav, Praha, SNTL, 1990.

[STR] Strejc, V.: Teorie lineární regulace, ČVUT, Praha, 1970.

[STR1] Strejc, V.: Teorie automatického řízení, ČVUT, Praha, 1982.

[UNB] Unbehauen, Heinz : Identification of continuous systems, New York, Elsevier Science, 1987.

[WAL] Walter Eric – Pronzato, L.: Identification of Parametric models from experimental data, Berlin, Springer, 1997.

[WIL] Willems, J. C.: From data to model, New York, Springer-Verlag, 1989.

[ZÍT] Zítek Pavel: Simulace dynamických systémů, Praha, SNTL, 1990.

REJSTŘÍK

A

AR model, 26, 28
ARMA model, 26, 28
ARMAX model, 26, 27
ARX model, 26, 27
autokovarianční
funkce, 49

B

bílý šum, 53
black box, 16, 20, 21, 26
Box-Jenkins
model, 26, 27

C

Coulombovo
tření, 78, 107, 121, 127

D

derivační
metoda, 58, 98, 112, 12
2, 123, 135
deterministická
optimalizace, 47
deterministické
modely, 25
deterministické
systémy, 11, 142
diferenciální
rovnice, 55, 62, 64, 66,
68, 70, 73, 76
diskretizace, 52
diskrétní
modely, 10, 24, 52
drift, 85

E

estimovaný
parametr, 49
exponenciální
zapomínání, 36

F

frekvenční
charakteristika, 49
fyzikální modely, 20

G

Galileo, 31
globální
minimum, 47, 55
globální
optimalizace, 47
gradientní metoda, 38
gradientní metody, 38

H

hustota
pravděpodobnosti, 31

C

Choleského
dekompozice, 32

I

identifikace, 17, 19
identifikační cyklus, 17

integrační
metoda, 53, 98, 112, 12
3, 135

inverzní model, 15
Inverzní model, 15

J

Jacobiho
matice, 75, 77, 78

K

konvergence, 73
kritérium maximální
věrohodnosti, 31
kroskovarianční
funkce, 50
kvadratické účelová
funkce, 135
kvalita
modelu, 17, 49, 98, 104
, 127, 134
kvazinewtonská
metoda, 38

L

Lagrangeovy
rovnice, 20
Laplaceův
operátor, 8, 23, 53
lineární
modely, 11, 22, 23, 37,
70
lineární subsystém, 143
lokální optima, 29
lokální optimum, 29

M

matematicko-fyzikální
analýza, 13, 108
měřená porucha, 14
metoda konjugovaných
směrů, 38
metoda
kvazilinearizace, 72,
75, 143
metoda nejmenších
čtverců, 30, 31, 33, 36,
37, 38
modulační
funkce, 59, 61, 63, 69

N

nelineární systémy, 70
nenulové počáteční
podmínky, 64, 65, 67, 6
8, 142
numerický
algoritmus, 18

O

OE model, 26, 28

P

parametry
trendu, 85, 86, 142
porucha na vstupu, 84
porucha na
výstupu, 84, 85
pravdivý model, 15, 20
pravdivý systém, 69
prof. Strejc, 10, 52
přenosové matice, 23, 24

Q

QR
dekompozice, 32, 33, 5
5

R

rekurzivní metoda
nejmenších
čtverců, 36
residua, 49, 50, 100, 101,
103, 104, 109, 111, 112
, 116, 117, 125, 126, 13
0, 131, 137, 138, 140, 1
41
rozptyl, 48, 49, 104, 109,
126, 131
rozptyl parametrů, 49

S

simulace, 20, 24, 49
stochastické modely, 25
struktura
modelu, 17, 18, 25, 28
struktury modelů, 25, 26

S

šum, 14

T

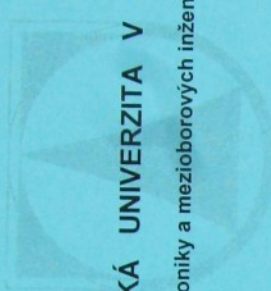
Taylorova řada, 72

U

účelová
funkce, 29, 30, 31, 32, 3
7, 38, 39, 40, 41, 43, 44,
47, 54, 55, 65, 72, 74, 7
5, 98, 112, 123, 135

V

vážená metoda
nejmenších
čtverců, 30
vliv poruch, 11, 84
vzorkovací
perioda, 10, 24, 25, 52,
86



TECHNICKÁ UNIVERZITA V LIBERCI

Fakulta mechatroniky a mezipřoborových inženýrských studií

Ing. Radomír Kamler

**METODY IDENTIFIKACE SPOJITÝCH MODELŮ
TECHNOLOGICKÝCH SOUSTAV**

Teze k disertační práci

Liberec 1999

TECHNICKÁ UNIVERZITA V LIBERCI

Fakulta mechatroniky a mezioborových inženýrských studií

Obor 26 – 15 – 9 „Technická kybernetika“

Katedra řídicí techniky

**METODY IDENTIFIKACE SPOJITÝCH MODELŮ TECHNOLOGICKÝCH
SOUSTAV**

Ing. Radomír Kamler

Teze k disertační práci

Liberec 1999

Disertační práce byla vypracována v rámci interní formy postgraduálního studia na katedře řídicí techniky, fakulty mechatroniky a meziporoborových inženýrských studií Technické Univerzity v Liberci.

Disertant: Ing. Radomir Kamlar
Katedra řídicí techniky
Fakulta mechatroniky a meziporoborových inženýrských studií
Technická Univerzita v Liberci

Školitel: Doc. Ing. Bedřich Janeček, CSc.
Katedra řídicí techniky
Fakulta mechatroniky a meziporoborových inženýrských studií
Technická Univerzita v Liberci

Oponenti: Doc. Ing. Jan Soukup, CSc.
Ústav přístrojové a řídicí techniky
Fakulta strojní
ČVUT Praha

Prof. Ing. Balátě, DrSc.
Katedra automatizace a řídicí techniky
Fakulta technologická Zlín
VUT Brno

Doc. Ing. Štefan Kozák, CSc.
Katedra automatizovaných systémů řízení
Fakulta elektroniky a informatiky
Slovenská technická univerzita v Bratislavě

Obhajoba disertační práce se koná dne: 27. 3. 2000

S disertační prací je možno se seznámit na děkanátě fakulty mechatroniky a meziporoborových inženýrských studií Technické Univerzity v Liberci, Hálkova 6, Liberec.

OBSAH

1	ÚVODNÍ POZNÁMKA	7
1.1	CÍLE DISERTAČNÍ PRÁCE	8
2	IDENTIFIKACE VE SPOJITÉ ČASOVÉ OBLASTI	9
2.1	INTEGRAČNÍ METODA	10
2.2	DERIVAČNÍ METODA	12
2.3	VLIV POČÁTEČNÍCH PODMÍNEK	16
2.3.1	APLIKACE NENULOVÝCH POČÁTEČNÍCH PODMÍNEK NA INTEGRAČNÍ METODU	17
2.3.2	APLIKACE NENULOVÝCH POČÁTEČNÍCH PODMÍNEK NA DERIVAČNÍ METODU	18
2.4	INTEGRAČNÍ A DERIVAČNÍ METODA PRO Nelineární systémy	19
2.5	METODA KVAZILINEARIZACE	20
2.6	KOMBINACE KVAZILINEARIZACE A INTEGRAČNÍ METODY	21
2.7	IDENTIFIKACE NEDIFERENCovatELNÝCH Nelineárních modelů	22
2.8	ROZBOR Vlivu poruch na odhad parametrů	25
2.9	Závěr	26
2.10	IDENTIFIKACE VÝŠKY HLADINY	27
3	Závěr	37

a_x, b_x	koefficienty přenosu
A, B, C, D	matice stavového popisu
$A(s), B(s)$	polynomy přenosu
$\arg \min J(p)$	hodnota p , která minimalizuje J
$E\{p\}$	střední hodnota vektoru p
$e(t)$	residuální odchylka
$\Phi(t)$	modulační funkce
I	jednotková matice
J	kvadratická účelová funkce
$n(t)$	šum
$n_m(t)$	měřená porucha
n_a, n_b	stupeň polynomů A, B
p	vektor hledaných parametrů
p_{ai}	vektor nelineárních parametrů
p_k	vektor p v kroku k
q	úroveň šumu
q^{-1}	operátor zpoždění
R	regresní matice
s	Laplaceův operátor
t	nezávislá proměnná, často čas
T	vzorkovací perioda
$u(t)$	vstupní veličina
w	váhový koefficient
$y(t)$	výstupní veličina
$y_m(t)$	výstup modelu
$Y(s)$	Laplaceův obraz funkce $y(t)$
$\langle x, y \rangle$	skalární součin vektorů x a y
$\dot{y}, y', y^{(i)}$	$\frac{dy}{dt}$
$\int_0^T y(t) dt$	$\int_0^T \int_0^T y(t) dt_j dt$
z	operátor z-transformace

ABSTRACT

This dissertation addresses one of the most challenging problems in system identification, the model estimation in continuous time domain. The material is presented in three main parts. The first one reviews some basic modeling and identification concepts, algorithms and ideas from which the following parts take off. An overview of different possible model structures is given in a common framework.

The second part is focused on new identification methods in continuous time domain. In this work we prefer that the model structure is chosen by semi-physical modeling. Efficient identification algorithms and procedures for these structures are developed and analyzed. These model structures are divided into three groups:

- models linear in parameters
- nonlinear models with differentiable cost function
- nonlinear models with non-differentiable cost function

The two methods (integrational and derivational) are used for the first group of models. These methods utilize least square estimator. It is suggested to extend estimation procedure by some iterative method in order to find better value of a cost function. Further, kvazilinearization as method for the second group of models, is investigated. A combination of kvazilinearization and integrational methods was tested for partially linear in parameter model structures. This makes it possible to significantly reduce the dimension of the search space in the most application of mechatronics systems and consequently speed up search for optimal parameters. General optimization methods are used for the third group of models.

The usefulness of the proposed methods is demonstrated through a number of concrete simulation and application studies. The robustness against different kinds of perturbations is also examined.

Finally and quite importantly, the third part concerns real identification applications and analyzes usefulness of proposed methods. With these methods, high quality linear and non-linear models are estimated in continuous time domain.

1 Úvodní poznámka

Výzkumná činnost v oblasti automatického řízení se v několika posledních desetiletích zaměřila především na metody identifikace v diskrétní časové oblasti. Metody pracující s modely ve spojitě časové oblasti se tak dostaly do stínu pozornosti. Přesto byly vyvinuty metody umožňující identifikaci ve spojitě časové oblasti, ale jejich použití je často spojeno se značnými omezeními (viz. metoda prof. Strejce s omezením na přechodové charakteristiky). S masivním rozvojem digitálních počítačů (v posledních 3 až 5 desetiletích) se zdá být logické, že byly vyvinuty metody hledání modelů pracující s daty v diskrétní formě (tedy stejně jak jsou prezentována data v digitálních počítačích). Tato podobnost umožnila vyvinout numericky velice efektivní algoritmy. Současně ale se zvyšováním výkonnosti komerčně využívaných počítačů můžeme dosáhnout vzorkovacích period tak malých (vzorkovací perioda několikanásobně menší než nejmenší časová konstanta identifikované technologie), že diskrétně nasnímaná data mohou nahradit spojitý signál. A společně s metodami identifikace ve spojitě časové oblasti můžeme těžit z výhod spojitých modelů v porovnání s diskrétními: u diskrétních modelů jsme omezení vzorkovací periodou, v jejichž násobcích je počítán výstup modelu. S tím souvisí, že diskrétní systémy v sobě mohou skrývat oscilace korespondujícího spojitého systému. A konečně parametry diskrétních systémů nemívají konkrétní význam a závisí na vzorkovací periodě. Nutno ještě podotknout, že vlastnosti spojitých a diskrétních systémů nejsou vždy analogické. V diskrétní oblasti můžeme např. realizovat oscilující soustavu prvního řádu, zatímco ve spojitě oblasti to možné není. Z toho také vyplývá obtížnost plnohodnotné transformace diskrétního popisu na spojitý. Tyto důvody nás vedly k tomu, abychom se v této práci zabývali metodami hledání modelů ve spojitě časové oblasti.

1.1 CÍLE DISERTAČNÍ PRÁCE

Cílem této práce je vyvinout a vyzkoušet metody identifikace deterministických lineárních a nelineárních modelů ve spojitě časové oblasti. V první části je stručně shrnut postup experimentální identifikace společný identifikaci jak v diskrétní tak i ve spojitě časové oblasti a popsán současný stav poznání týkající se identifikace ve spojitě časové oblasti. V dalších částech se práce zabývá způsoby, jak základní identifikační objekty a nástroje (data, modely, optimalizace) zapadají do procesu hledání vhodného modelu. Tato úvodní část práce definuje nástroje pro následné zkoumání možnosti identifikace ve spojitě časové oblasti.

Vlastní těžiště práce spočívá v metodách hledání modelů ve spojitě časové oblasti. Ty se dají rozdělit do tří skupin dle charakteru modelu:

- modely lineární v parametrech
- modely nelineární, u kterých známe parciální derivace výstupu dle vektoru hledaných parametrů
- modely nelineární, u kterých neznáme parciální derivace výstupu dle vektoru hledaných parametrů

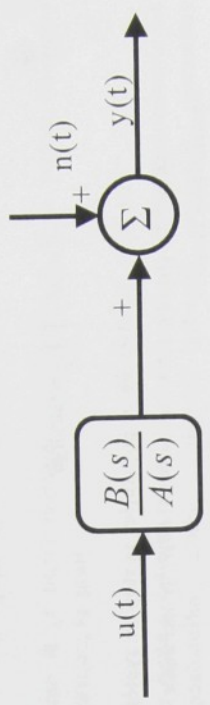
V práci je také stručně analyzován vliv poruch na kvalitu nalezeného modelu. Závěrečná část je věnována praktickým ukázkám a jednotlivé metody jsou vyzkoušeny na několika skutečných systémech a nalezené modely jsou podrobeny kvalitativní analýze.

2 Identifikace ve spojité časové oblasti

V historii vývoje identifikačních metod se spojité časové oblasti v posledních desetiletích věnovalo velmi málo úsilí a v literatuře jsou jen spíše popsány některé metody hledání spojitých modelů. Ačkoliv počátky vývoje teorie identifikace jsou spojeny se spojitými modely (připomeňme metodu prof. Strejce), tyto byly velmi brzy nahrazeny modely diskrétními, které pracují s daty tak jak jsou dnes obvykle snímána tzn. v diskrétní formě. Tyto metody zaznamenaly v poslední době ohromný rozmach a zcela zastínily algoritmy hledající modely ve spojité časové oblasti. Naproti tomu, s rostoucí rychlostí zpracování dat současnými počítači jsme schopni snímat data s tak malou vzorkovací periodou, že takto můžeme nahradit spojitý signál pro většinu mechatronických technologií (za předpokladu, že vzorkovací perioda je několikanásobně menší než nejvyšší časová konstanta systému). A i když zpracování takového velkého množství dat klade nároky na výkon počítače (popř. na čas výpočtu) můžeme zároveň těžit z výhod, které nám poskytne model systému ve spojité časové oblasti. Zmíňme se alespoň o některých v porovnání s modely v diskrétní oblasti. Tyto výhody vycházejí z jednoduchého principu: studovaný systém se obvykle nachází ve spojité časové oblasti a proto je výhodné ho i tak popisovat. Naproti tomu výstupy diskrétních modelů jsou počítány pouze v určitých časových okamžicích odpovídajících vzorkování a tím nás trochu svazují. Z toho plyne, že výstup diskrétního modelu v sobě může skrývat oscilace odpovídajícího spojitého modelu. A konečně parametry diskrétních modelů nemají všeobecně jasný fyzikální význam (parametry získané diskretizací závisí na volbě vzorkovací periody) a ani vlastnosti diskrétních a spojitých modelů nejsou vždy analogické. Z toho také vyplývá, že sestavování spojitých modelů z diskrétních může být složité a nepřesné.

2.1 INTEGRAČNÍ METODA

Pro jednoduchost zpočátku uvažujeme jednorozměrný systém, jehož výstup lze popsat lineární diferenciální rovnicí v závislosti na vstupu a vnitřních proměnných. Dále připouštíme, že na výstup ze systému může působit porucha, kterou zase pro jednoduchost nejprve uvažujeme v podobě bílého šumu (obrázek 2.1-1):



obrázek 2.1-1: Lineární jednorozměrný model

Známo: $u(t), y(t)$.
Neznámo: $B(s), A(s), n(t)$.
Zjišťujeme: $B(s), A(s)$,

kde s ...je Laplaceův operátor,

na ...je stupeň polynomu $A(s)$ viz. (2-1),

nb ...je stupeň polynomu $B(s)$ viz. (2-2).

$$A(s) = a_0 + a_1s + a_2s^2 + \dots + a_{na-1}s^{na-1} + s^{na}, \quad (2-1)$$

$$B(s) = b_0 + b_1s + b_2s^2 + \dots + a_{nb}s^{nb}. \quad (2-2)$$

Diferenciální rovnici popisující chování takového systému lze psát ve tvaru:

$$y^{(na)}(t) + \sum_{i=0}^{na-1} a_i y^{(i)}(t) = \sum_{j=0}^{nb} b_j u^{(j)}(t) \quad (2-3)$$

a tu pak jednoduše upravíme na (2-4):

$$(2-4)$$

$$y(t)^{(m)} = -\sum_{i=0}^{m-1} a_i y^{(i)}(t) + \sum_{i=0}^{m-1} b_i u^{(i)}(t)$$

a po na integracích (2-4) dostaneme (2-5):

$$y(t) = -\sum_{i=0}^{m-1} a_i \int_0^t y(t)dt + \sum_{i=0}^{m-1} b_i \int_0^t u(t)dt, \tag{2-5}$$

kde $\int_0^t$ je označení i-tých integrálů, takže např.

$$\int_0^t y(t)dt = \int_0^t \left(\int_0^t y(t)dt \right) dt.$$

Tim jsme získali rovnici popisující výstup ze systému lineární v hledaných parametrech a_i a b_i . Tudiž ji můžeme přepsat v maticovém tvaru:

$$y_m = Xp, \tag{2-6}$$

kde

$$X = \left[\int_0^t u dt, \int_0^t u dt \dots, \int_0^t u dt, -\int_0^t y dt, -\int_0^t y dt, \dots, -\int_0^t y dt \right],$$

$$p = [b_0, b_1, \dots, b_{m-1}, a_0, a_1, \dots, a_{m-1}]^T, \int_0^t y dt = \left[\int_0^t y(t) dt \int_0^{2T} y(t) dt \dots \int_0^{NT} y(t) dt \right]^T.$$

Jestliže pro nalezení optimálních parametrů ze změřených signálů $u(t)$ a $y(t)$ použijeme metodu menších čtverců, kde minimalizovaná účelová funkce je definována jako kvadrát rozdílu odchylek měření a modelu viz. (2-7):

$$J = e^T e \rightarrow \min, \tag{2-7}$$

kde $e = y - y_m$.

Minimalizací účelové funkce (2-7) získáme vztah (2-8) pro optimální parametry:

$$\frac{\partial J}{\partial p} = -\frac{1}{2} X^T (y - X\hat{p}) = 0 \tag{2-8}$$

a za předpokladu, že existuje inverze matice $X^T X$ můžeme psát vektor parametrů následovně (2-9):

$$\hat{p} = (X^T X)^{-1} X^T y \tag{2-9}$$

Jak již bylo naznačeno v předchozích kapitolách, v praxi se model většinou nehledá pomocí inverze matice $X^T X$ kvůli špatným numerickým vlastnostem této metody. Proto se velmi často používají metody dekompozice matice X (jmenujeme alespoň Choleského dekompozici, QR dekompozici... [PRE], [NIU]).

Pro tento případ můžeme QR dekompozici popsat následovně:

- matici X rozložíme tak, že:

$$X = Q * [R \ 0]^T \text{ a } Q^T y = [z_1 \ z_2]^T,$$

kde R je horní trojúhelníková matice a Q je taková čtvercová matice, že platí:

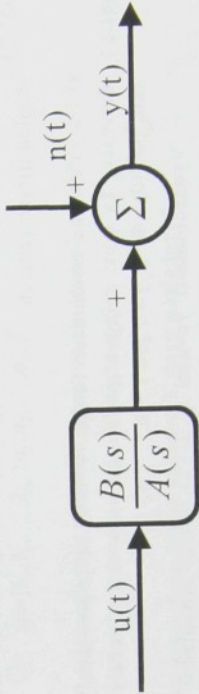
$$Q^T \cdot Q = E$$

a poté vyřešením soustavy rovnic $R\hat{p} = z_1$ prostým dosazením od posledního řádku nahoru získáme optimální parametry identifikovaného modelu.

Pozn. Nutno podotknout, že takto nalezené parametry nemusí být optimální z hlediska globálního minima účelové funkce, protože ve výpočtu nahrazujeme vnitřní proměnné modelu změřenými. To se nám negativně projeví hlavně pokud se tyto od sebe podstatně liší (např. při špatné volbě struktury modelu). I přes to ale lze tuto metodu doporučit pro její numerickou nenáročnost a dosažené výsledky je možné např. použít jako inicializační odhad následujících iteračních optimalizačních algoritmů.

2.2 DERIVAČNÍ METODA

Uvažujme znovu lineární jednorozměrný model jako tomu bylo u integrační metody. Tento model lze graficky vyjádřit na následujícím grafu (obrázek 2.2-1):



obrázek 2.2-1: Lineární jednorozměrný model

Známo: $u(t), y(t)$.

Neznámo: $B(s), A(s), n(t)$.

Zjišťujeme: $B(s), A(s)$.

Pro změnu, bez újmy na obecnosti, budeme uvažovat polynom $A(s)$ v trošku odlišném tvaru od (2-1), $B(s)$ zůstává stejný:

$$A(s) = 1 + a_1 s + a_2 s^2 + \dots + a_{n-1} s^{n-1} + a_n s^n, \quad (2-10)$$

$$B(s) = b_0 + b_1 s + b_2 s^2 + \dots + a_{nb} s^{nb}. \quad (2-11)$$

Diferenciální rovnici popisující chování takového systému lze psát ve tvaru:

$$y(t) + \sum_{i=1}^{na} a_i y^{(i)}(t) = \sum_{i=0}^{nb} b_i u^{(i)}(t) \quad (2-12)$$

a to stejně jako v předchozí kapitole 2.1 přepíšeme na:

$$y(t) = \sum_{i=0}^{nb} b_i u^{(i)}(t) - \sum_{i=1}^{na} a_i y^{(i)}(t), \quad (2-13)$$

čímž jsme dostali rovnici popisující výstup ze systému v závislosti na derivacích vstupu a výstupu. Samozřejmě by bylo možné tyto derivace numericky vypočítat a přímo použít v modelu. To je ale ve většině případů krajně nevhodné, protože vysokofrekvenční šum, který nám ve většině případů zhoršuje měření, by se po několikanásobné derivaci drasticky zvětšil a tím nám zastínil užitečné informace o systému. Proto použijeme přístup tzv. *modulačních funkcí*, jehož princip se pokusíme objasnit na následujících řádcích.

Vyberme modulační funkci Φ tak, že známe analyticky její derivace. Ty musí splňovat podmínku nulovosti v počátku a konci, tzn. $\Phi^{(i)}(0) = \Phi^{(i)}(t_f) = 0$ pro $i = 0, \dots, n_a - 1$.

Vynásobíme-li rovnici (2-13) touto modulační funkcí a provedeme-li integraci dostaneme (2-14):

$$\int_0^{t_f} y(t) \Phi(t) dt = \sum_{i=0}^{nb} b_i \int_0^{t_f} u^{(i)}(t) \Phi(t) dt - \sum_{i=1}^{na} a_i \int_0^{t_f} y^{(i)}(t) \Phi(t) dt. \quad (2-14)$$

Rovnici (2-14) lze poté za předpokladů o modulační funkci přepsat na rovnici (2-15):

$$\int_0^{t_f} y(t) \Phi(t) dt = \sum_{i=0}^{nb} b_i (-1)^i \int_0^{t_f} u(t) \Phi^{(i)}(t) dt - \sum_{i=1}^{na} a_i (-1)^i \int_0^{t_f} y(t) \Phi^{(i)}(t) dt. \quad (2-15)$$

A zde vidíme cíl celého snažení: v rovnici (2-15) jsme se zbavili derivací měřeného signálu. Ty jsou nahrazeny analyticky známými derivacemi modulační funkce.

Protože data jsou ve skutečnosti snímána diskrétně, lze rovnici (10) přepsat ve vektorovém tvaru:

$$\langle y | \Phi \rangle = \sum_{i=0}^{nb} b_i \langle (-1)^i \langle u | \Phi^{(i)} \rangle \rangle - \sum_{i=1}^{na} a_i \langle (-1)^i \langle y | \Phi^{(i)} \rangle \rangle, \quad (2-16)$$

kde $\langle y | \Phi \rangle$ značí skalární součin.

Nevýhodou této metody je, že volba modulační funkce není triviální a pro zjištění parametrů modelu potřebujeme nejméně tolik nezávislých testovacích funkcí, kolik máme neznámých parametrů modelu. Jeden z možných postupů sestavení modulační funkce si předvedeme v následujícím odstavci.

Pro potřeby minimalizace metodou nejmenších čtverců sestavíme matici R takto:

$$X(k,:) = \left[\langle u | \Phi_k \rangle, -\langle u | \Phi_k \rangle, \langle u | \Phi_k \rangle, \dots, (-1)^{nb} \langle u | \Phi_k^{(nb)} \rangle, \right. \\ \left. \langle y | \Phi_k \rangle, -\langle y | \Phi_k \rangle, \langle y | \Phi_k \rangle, \dots, (-1)^{na} \langle y | \Phi_k^{(na)} \rangle \right], \quad (2-17)$$

kde: $X(k,:)$ značí k -tý řádek matice X ,

$\Phi_k \dots k$ -tá lineárně nezávislá modulační funkce.

Vektor hledaných parametrů p sestavíme dle (2-18):

$$p = [b_0, b_1, \dots, b_{nb}, a_1, a_2, \dots, a_{na}]^T. \quad (2-18)$$

A minimalizuje metodou nejmenších čtverců rovnici (2-19):

$$Y = Xp, \quad (2-19)$$

ve které jsou řádky vektoru Y dány následovně:

$$y(k) = \langle y | \Phi_k \rangle, \text{ kde } y(k) \dots \text{ je } k\text{-tý prvek vektoru } Y.$$

Dále již pak při minimalizaci postupujeme standardními kroky metody nejmenších čtverců stejně jako u integrační metody a nalezneme vektor optimálních parametrů lineárního modelu.

2.3 Vliv počátečních podmínek

V praxi se mnohdy setkáváme s technologiemi, u kterých nejsme schopni ovlivnit vstupní signál natolik, abychom mohli započít měření z nulových počátečních podmínek. Situace se dále komplikuje v případě, když je v systému ještě navíc akumulována energie, tzn. že ani vnitřní stavové veličiny nejsou nulové na počátku měření. S nenulovou počáteční podmínkou na výstupu systému jsme schopni se velkou snadno vyrovnat (např. vyloučením tzv. driftu [OLE]). Horší je tomu s nenulovými počátečními podmínkami vyšších řádů(např. u pohybujícího se tělesa s rychlostí a zrychlením). Při dostatečně dlouhém souboru dat se dají zanedbat, zvláště když použijeme iterační metodu s exponenciálním zapomináním, ale i tak ovlivňují odhad parametrů soustavy. V této kapitole bude předveden postup, jak se lze s tímto problémem vyrovnat.

Uvažujme znovu lineární systém popsaný diferenciální rovnici ve tvaru (2-20):

y^{(na)}(t) + \sum_{i=0}^{na-1} a_i y^{(i)}(t) = \sum_{i=0}^{mb} b_i u^{(i)}(t), (2-20)

ale tentokrát s nenulovými počátečními podmínkami y^{(i)}(0) \neq 0, i = 0, 1, \dots, na-1.

Řešení výstupu takového systému s nulovými počátečními podmínkami v Laplaceově obraze lze psát:

Y(s) = \frac{b_{mb}s^{mb} + b_{mb-1}s^{mb-1} + \dots + b_0}{s^{na} + a_{na-1}s^{na-1} + a_{na-2}s^{na-2} + \dots + a_0} U(s) = \frac{B(s)}{A(s)} U(s). (2-21)

Dle principu superpozice lze pro lineární systémy rozložit řešení s nenulovými počátečními podmínkami na řešení soustavy s nulovými počátečními podmínkami buzené vstupním signálem u(t) a řešení soustavy buzené pouze počátečními podmínkami:

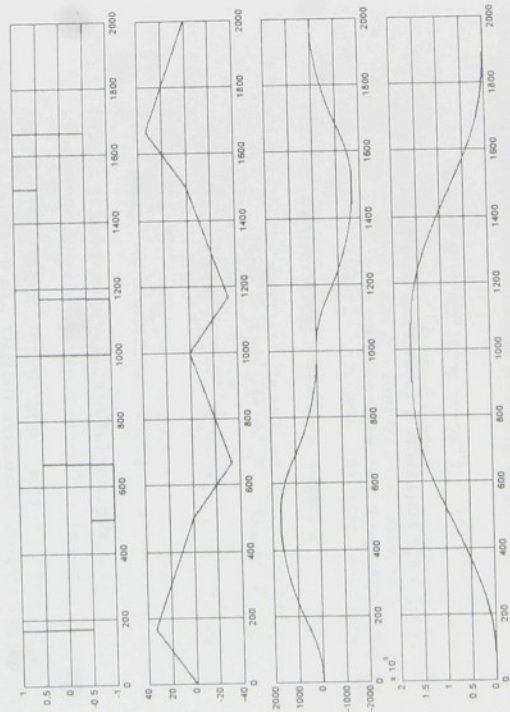
Y(s) = \frac{b_{mb}s^{mb} + b_{mb-1}s^{mb-1} + \dots + b_0}{s^{na} + a_{na-1}s^{na-1} + a_{na-2}s^{na-2} + \dots + a_0} \frac{U(s)}{s} + \frac{s^{na}y(0) + s^{na-1}[a_{na-1}y(0) + a_2y'(0) + \dots + a_{na-1}y^{(na-1)}(0)] \frac{1}{s}}{A(s)}

Po stručné obecné úvaze se podrobněji zabývájme tím, jak se nenulové počáteční podmínky projeví v integračním a derivačním identifikačním algoritmu (viz kapitoly 2.3.1 a 2.3.2).

Generování lineárně nezávislých modulačních funkcí:

Zmíňme se o jednom možném způsobu vytváření modulačních funkcí. Při generování modulačních funkcí postupujeme v opačném sledu, tzn. že nejdříve vytvoříme nejvyšší derivaci této funkce:

-ta je po částech konstantní a skládá se z 2^{na} úseků. Tyto úseky jsou generovány jako řada vznikající zrcadlovým otočením předcházející posloupnosti. Základem řady jsou dva úseky a a -b jejichž integrál je nulový. Např. pro na=3 se bude nejvyšší derivace modulační funkce skládat z osmi konstantních úseků: {a -b | -a b | -b a b a -b} (viz. obrázek 2.2-2 nahoře). Nižší derivace následovně získáme postupnou numerickou integrací. Modulační funkce pro soustavu třetího řádu jsou znázorněny v následujícím grafu (viz. obrázek 2.2-2).



obrázek 2.2-2: Modulační funkce a její derivace (ze zdola: \phi(t), \phi'(t), \phi''(t), \phi'''(t))

2.3.1 Aplikace nenulových počátečních podmínek na integrační metodu

Přepíšeme-li řešení jednorozměrného systému s nenulovými počátečními podmínkami a upravíme-li ho pro integrační metodu dostaneme (2-22):

y(t) = \sum_{i=0}^{mb} b_i \int_0^t u(t) dt - \sum_{i=1}^{na} a_i \int_0^t y(t) dt + \sum_{i=0}^{na-1} k_i \int_0^t \eta(t) dt, \tag{2-22}

kde

k_0 = y(0)
k_1 = a_{na-1} y(0) + y'(0)
k_2 = a_{na-2} y'(0) + a_{na-1} y''(0) + y''(0)
k_3 = a_{na-3} y''(0) + a_{na-2} y'''(0) + a_{na-1} y^{(4)}(0) + y^{(4)}(0)
:
k_{na-1} = a_1 y(0) + a_2 y'(0) + ... + a_{na-1} y^{(na-2)}(0) + y^{(na-1)}(0),

což můžeme znovu přepsat v maticovém tvaru:

y = Xp, \tag{2-23}

X = [\int_0^t u(t) dt, \int_0^t u(t) dt \dots, \int_0^t u(t) dt | - \int_0^t y(t) dt, - \int_0^t y(t) dt, \dots, \int_0^t y(t) dt, \dots, - \int_0^t y(t) dt | \eta(t), \int_0^t \eta(t), \dots, \int_0^t \eta(t)],

p = [b_0, b_1, \dots, b_{mb} | a_1, a_2, \dots, a_{na} | k_0, k_1, \dots, k_{na-1}]^T.

Zahrnutím nenulových počátečních podmínek do struktury modelu se vektor hledaných parametrů rozšířil o na koeficientů k_i. Z těchto koeficientů jsme případně schopni vypočítat počáteční podmínky. Ty nás ale ve většině případů nezajímají. Zvětšení počtu parametrů u metody nejmenších čtverců, pro kterou známe analytickou formuli pro optimální parametry, nehraje z numerického tak velkou roli a neklade přehnané nároky na dobu výpočtu. Jinak je tomu u obecných optimalizačních metod, kde každé zvýšení dimenze prohledávaného prostoru značně ztěžuje hledání optima účelové funkce [MAN].

2.3.2 Aplikace nenulových počátečních podmínek na derivační metodu

Přepíšeme-li řešení jednorozměrného systému s nenulovými počátečními podmínkami a upravíme-li ho pro derivační metodu dostaneme (2-24):

\int_0^t y(t) \Phi(t) dt = \sum_{i=0}^{mb} b_i (-1)^i \int_0^t u(t) \Phi^{(i)}(t) dt - \sum_{i=1}^{na} a_i (-1)^i \int_0^t y(t) \Phi^{(i)}(t) dt + \sum_{i=0}^{na-1} k_i \int_0^t \eta(t) \Phi^{(na-i)}(t) dt. \tag{2-24}

A protože testovací (modulační) funkce byly definovány tak, že platí:

\int_0^t \Phi^{(i)}(t) dt = 0 \text{ pro } i = 1, 2, \dots, na, \text{ pak i: } k_i \int_0^t \eta(t) \Phi^{(i)}(t) dt = 0,

což dosadíme-li do (2-24) dostaneme vztah (2-25), který je stejný jako tomu bylo bez počátečních podmínek viz. (2-15).

\int_0^t y(t) \Phi(t) dt = \sum_{i=0}^{mb} b_i (-1)^i \int_0^t u(t) \Phi^{(i)}(t) dt - \sum_{i=1}^{na} a_i (-1)^i \int_0^t y(t) \Phi^{(i)}(t) dt, \tag{2-25}

který nezávisí na počátečních podmínkách. Tzn., že můžeme derivační metodu používat nezávisle na tom, jestli byl systém na počátku měření s nenulovými počátečními podmínkami či nikoli. To je samozřejmě velká výhoda derivační metody oproti integrační, kde se počet parametrů při nenulových počátečních podmínkách zvyšoval v závislosti na řádu modelu.

2.5 METODA KVAZILINEARIZACE

Přejdeme k jinému přístupu identifikace a zabýváme se soustavami, jejichž výstup závisí alespoň na jednom nelineárním hledaném parametru. V tomto případě budeme řešit odhad optimálních parametrů systému z měřených dat tzv. metodou kvazilinearizace.

Předpokládejme, že výstup ze soustavy $y(t)$ je až na vektor hledaných parametrů $\mathbf{p}$ znám a je dán funkcí:

$$y(t) = f(\mathbf{p}, t). \quad (2-28)$$

Dále předpokládejme, že jsme pro studovaný systém zvolili nelineární model ve tvaru:

$$\mathbf{y}_m = f(\mathbf{p}, t). \quad (2-29)$$

Účelovou funkci definujeme jako součet kvadrátů odchylek modelu od systému:

$$J = \mathbf{e}^T \mathbf{e}, \quad (2-30)$$

$$\mathbf{e} = \mathbf{y} - \mathbf{y}_m = \mathbf{y} - f(\mathbf{p}, t). \quad (2-31)$$

Protože výstup modelu je nelineární funkcí neznámých parametrů, je třeba ho před použitím lineárního estimátoru nejprve linearizovat pomocí Taylorova rozvoje (uvážujeme pouze první dva členy, zbytek zanedbejme):

$$\mathbf{y}_m = f(\mathbf{p}, t) \approx \mathbf{y}_m + \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \Delta \mathbf{p}. \quad (2-32)$$

Minimalizací účelové funkce dané (2-30) dostaneme podmínku optimálního kroku:

$$\frac{\partial J}{\partial \mathbf{p}} = 2 \frac{\partial \mathbf{e}^T}{\partial \mathbf{p}} \mathbf{e} = 0, \quad (2-33)$$

a dosazením (2-32) do (2-31) dostáváme:

$$\mathbf{e} = \mathbf{y} - \mathbf{y}_m = \mathbf{y} - \left(\mathbf{y}_m + \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \Delta \mathbf{p} \right). \quad (2-34)$$

Po dosazení (2-34) do (2-33) získáme:

$$2 \frac{\partial \mathbf{y}_m^T}{\partial \mathbf{p}} \left[\mathbf{y} - \left(\mathbf{y}_m + \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \Delta \mathbf{p} \right) \right] = 0, \quad (2-35)$$

2.4 INTEGRAČNÍ A DERIVAČNÍ METODA PRO NELINEÁRNÍ SYSTÉMY

Nasnadě je jednoduché rozšíření výše popsaných identifikačních metod na nelineární modely, které ale jsou lineární v parametrech. Pro takto definované struktury modelů lze použít metodu nejmenších čtverců společně s derivační popř. integrační metodou. Vyjádříme jednorozměrný model ve tvaru (2-26):

$$\frac{d^m}{dt^m} y(t) + \sum_{i=0}^{m-1} a_i \frac{d^i}{dt^i} y(t) = \sum_{i=0}^{n-1} b_i \frac{d^i}{dt^i} u(t) + \sum_{i=1}^m p_i z_i(t, u, y), \quad (2-26)$$

kde $z_i(t, u, y)$ jsou nelineární funkce, o kterých předpokládáme, že na nich model závisí lineárně (přes hledané koeficienty). Příkladem takové nelineární funkce může být např. kvadrát ze vstupního signálu u^2 , odmocnina z tohoto signálu $u^{1/2}$, součin vstupního a výstupního signálu $y \cdot u^2$, atd.

Pro *integrační metodu* definujeme výstup modelu v maticovém tvaru (2-27):

$$\mathbf{y}_m = \mathbf{X} \mathbf{p}, \quad (2-27)$$

kde

$$\mathbf{X} = \left[\int u(t) dt, \int u(t) dt, \dots, \int u(t) dt, - \int y(t) dt, - \int y(t) dt, \dots, - \int y(t) dt \right] \\ \int z_1(t) dt, \dots, \int z_m(t) dt, \\ \mathbf{p} = [b_0, b_1, \dots, b_{n-1} | a_0, a_1, \dots, a_{m-1} | p_1, \dots, p_m]^T.$$

Podobně můžeme sestavit model pro derivační metodu. Takto sestavené modely nám umožní použít metodu nejmenších čtverců a následně nalezení vektoru optimálních parametrů.

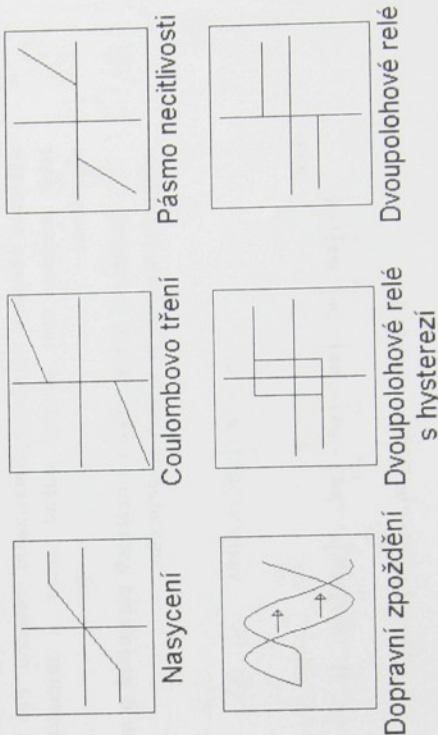
který lze také psát v následující podobě:

$$\mathbf{p}^{k+1} = \mathbf{p}^k + [\mathbf{J}_k^T \mathbf{J}_k]^{-1} \mathbf{J}_k^T (\mathbf{y} - \mathbf{y}_m^k), \text{ kde } \mathbf{J}_k^T \text{ je Jacobiho matice a je v našem případě dána z (2-39) a definována:}$$

$$\mathbf{J}_k = \left[\begin{matrix} \frac{\partial y_1}{\partial p_1} & \dots & \frac{\partial y_1}{\partial p_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial p_1} & \dots & \frac{\partial y_m}{\partial p_m} \end{matrix} \right] = \left[\int_{t_0}^{t_1} \mathbf{y}^T(t) \mathbf{A}_1(t) \mathbf{p} dt, \dots, \int_{t_0}^{t_1} \mathbf{y}^T(t) \mathbf{A}_m(t) \mathbf{p} dt \right]^T$$

2.7 IDENTIFIKACE NEDIFERENCOVATELNÝCH NELINEÁRNÍCH MODELŮ

V praxi se často setkáváme s nelineárními systémy, jejichž podstata neumožňuje lineární aproximaci. Mezi tyto tzv. "tvrdé nelinearity", které se často objevují v mechatronických systémech, patří Coulombovo tření, nasycení, pásmo necitlivosti, hystereze atd. (obrázek 2.7-1). Tato skupina nelinearit má společnou nepřijemnou vlastnost: nelze analyticky určit jejich derivace dle hledaných parametrů. A z toho vyplývající nepřijemným důsledkem je, že pro nalezení optimálního parametru modelu s tvrdou nelinearitou lze jen zřídka použít některou z dosud uvedených identifikačních metod.



obrázek 2.7-1: Typické příklady tvrdých nelinearit

$$\text{z něhož } \Delta \mathbf{p} = \left[\frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \right]^{-1} \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} (\mathbf{y} - \mathbf{y}_m). \quad (2-36)$$

(2-36) použijeme ve vztahu (2-37) pro úpravu vektoru parametrů iteracním výpočtem :

$$\mathbf{p}^{k+1} = \mathbf{p}^k + \left[\frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \right]^{-1} \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} (\mathbf{y} - \mathbf{y}_m^k) \quad (2-37)$$

V případě, že se při iteracním výpočtu vektor parametrů již dále nemění, výpočet ukončíme (dostáváme tzv. bodový odhad parametrů). Nevýhodou této metody je nutnost počátečního odhadu vektoru $\mathbf{p}$ a také to, že konvergence není zaručena. Podmínky konvergence lze vymezit analyticky [SOU].

2.6 KOMBINACE KVAZILINEARIZACE A INTEGRAČNÍ METODY

Pokusme se zkombinovat integrační metodu a metodu kvazilinearizace a tak těžit z výhod obou u částečně lineárních struktur modelů. Výhodou kvazilinearizace je schopnost pracovat s nelineárními modely, její nevýhodou je vyšší numerická náročnost a nutnost inicializačního odhadu parametrů. Naopak integrační metoda je rychlá a numericky ne příliš náročná, zato ale nedokáže pracovat s modely nelineárními v parametrech. Uvažujme nelineární systém ve tvaru:

$$\mathbf{y}^{(m)}(t) + \sum_{i=0}^{m-1} a_i \mathbf{y}^{(i)}(t) = \sum_{j=0}^{n_b} b_j \mathbf{u}^{(j)}(t) + f(\mathbf{p}_m, \mathbf{u}^{(j)}), \quad (2-38)$$

kde $f(\mathbf{p}_m, \mathbf{u}^{(j)})$ je nelineární funkce hledaného vektoru parametrů $\mathbf{p}_m$ a obecně i-tých derivací vstupu a výstupu ze systému.

Provedeme-li úpravu (2-38) na tvar vhodný pro integrační metodu a označíme jako výstup modelu, dostaneme:

$$\mathbf{y}_m(t) = - \sum_{i=0}^{m-1} a_i \int_{t_0}^t \mathbf{y}^{(i)}(t) dt + \sum_{j=0}^{n_b} b_j \int_{t_0}^t \mathbf{u}^{(j)}(t) dt + \int_{t_0}^t f(\mathbf{p}_m, \mathbf{u}^{(j)}, \mathbf{y}^{(j)}). \quad (2-39)$$

Pro metodu kvazilinearizace jsme v kapitole 2.5 získali vztah:

$$\mathbf{p}^{k+1} = \mathbf{p}^k + \left[\frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} \right]^{-1} \frac{\partial \mathbf{y}_m}{\partial \mathbf{p}} (\mathbf{y} - \mathbf{y}_m^k), \quad (2-40)$$

Tento problém je proto nutné řešit obecnými optimalizačními metodami. Bohužel náročnost času výpočtu u obecných optimalizačních metod značně roste s počtem hledaných parametrů. V mnoha případech ale můžeme v modelu nalézt lineární část a tu pak řešit metodou nejmenších čtverců a tak snížit počet parametrů hledaných obecnou optimalizační metodou.

Uvažujme popis nelineárního systému ve tvaru (2-41):

$$y^{(n)}(t) + \sum_{i=0}^{n-1} a_i y^{(i)}(t) = \sum_{j=0}^m b_j u^{(j)}(t) + \sum_{l=1}^m p_l z_l(t, u, y) + f(p_{n1}, u, y, t), \quad (2-41)$$

kde první tři členy tohoto modelu tvoří lineární subsystém, čtvrtý člen $\sum_{l=1}^m p_l z_l(t, u, y)$ je lineární alespoň v parametrech, i když obsahuje nelineární funkce $z_l(t, u, y)$, a konečně poslední člen $f(p_{n1}, u, y, t)$ je nelineární funkce hledaného vektoru parametrů p_{n1} a obecně i vstupu a výstupu ze systému a času.

Vektor lineárních parametrů, které budeme hledat metodou nejmenších čtverců, vytvoříme následovně:

$$p = [b_0, b_1, \dots, b_{nb} | a_0, a_1, \dots, a_{na-1} | p_1, \dots, p_m]^T.$$

Zbytek hledaných parametrů, vektor p_{n1} budeme hledat některou obecnou iterační optimalizační metodou. Pro tento vektor také musíme znát prvotní odhad nutný k inicializaci celého algoritmu.

Budeme-li hledat vektor lineárních parametrů p např. pomocí integrační metody, je výstup modelu dán:

$$y_m = Xp + \int_{na}^s f(\hat{p}_{n1}^s, u, y, t) dt, \quad (2-42)$$

kde:

$$X = \left[\int_{na}^s u dt, \int_{na}^s u dt, \dots, \int_{na}^s u dt \right] - \left[\int_{na}^s y dt, \dots, \int_{na}^s y dt \right] \int_{na}^s z_1 dt, \dots, \int_{na}^s z_m dt,$$
$$p = [b_0, b_1, \dots, b_{nb} | a_0, a_1, \dots, a_{na-1} | p_1, \dots, p_m]^T.$$

Celý algoritmus se skládá z následujících kroků:

- $k=1$, volba inicializačního vektoru nelineárních parametrů $\hat{p}_{n1}^k$.
- Metodou nejmenších čtverců nalezneme vektor optimálních lineárních parametrů $\hat{p}^k$ pro vektor $\hat{p}_{n1}^k$.
- Obecnou optimalizační metodu nalezneme $\hat{p}_{n1}^{k+1}$. Při hledání dosadíme za ostatní parametry modelu vektor $\hat{p}^k$.
- Zvýšíme k o jedničku a opakujeme algoritmus od druhého kroku.
- Cyklus opakujeme tak dlouho, dokud se nám mění vektor hledaných parametrů (v určité toleranci), popř. dokud jsme nevyčerpali max. počet iterací.

Parametry nalezené tímto algoritmem je vhodné použít jako inicializační pro některou obecnou optimalizační metodu a tak se pokusit nalézt lepší model. Toto doporučení vyplývá z nectnosti integrační (resp. derivační) metody zmíněné v kapitole 2.9.

2.8 ROZBOR Vlivu PORUCH NA ODHAD PARAMETRŮ

Předpokládejme, že naměřená data byla vytvořena systémem s vektorem parametrů $\mathbf{p}^*$ a měření bylo zatíženo náhodnou sekvencí poruch $\mathbf{n}_0$:

$$\mathbf{y} = \mathbf{R}\mathbf{p}^* + \mathbf{n}_0. \quad (2-43)$$

Dle metody nejmenších čtverců je estimovaný odhad vektoru parametrů dán dle

$$\hat{\mathbf{p}}_L = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{y}. \quad (2-44)$$

Následovně dostaneme

$$\hat{\mathbf{p}}_L = \mathbf{p}^* + (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{n}_0. \quad (2-45)$$

Požadavkem je, aby vektor odhadu parametrů $\hat{\mathbf{p}}_L$ byl blízký skutečnému $\mathbf{p}^*$ a konvergoval k němu když se počet vzorků na systému blíží k nekonečnu. Vektor odhadu parametrů $\hat{\mathbf{p}}_L$ je nestraným odhadem, tzn. že jeho průměrná hodnota pro všechna data, která by mohla být získána na systému s parametry $\mathbf{p}^*$, konverguje ke skutečné hodnotě

$$E_{y/p} \{\hat{\mathbf{p}}_L\} = \mathbf{p}^*, \quad (2-46)$$

pouze a jen tehdy, jestliže

$$E_{y/p} \{(\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{n}_0\} = \mathbf{0}. \quad (2-47)$$

Tzn. proto aby estimovaný odhad parametrů $\hat{\mathbf{p}}_L$ konvergoval ke skutečnému $\mathbf{p}^*$ musí být splněno:

- matice $E\{\mathbf{R}^T \mathbf{R}\}$ není singulární a
- jestliže $E\{\mathbf{R}^T \mathbf{n}_0\} = \mathbf{0}$. Což bude splněno buď, když
 - $\{\mathbf{n}_0\}$ je sekvence nezávislých náhodných poruch s nulovou střední hodnotou (bílý šum). Potom $\{\mathbf{n}_0\}$ nezávisí na tom co se stalo v minulosti a
$$E\{\mathbf{R}^T \mathbf{n}_0\} = \mathbf{0}.$$

nebo

- vstupní signál $\{u(t)\}$ je nezávislý na náhodné poruše $\{\mathbf{n}_0\}$ s nulovou střední hodnotou a $\mathbf{R}$ neobsahuje členy $\mathbf{y}$.

2.9 ZÁVĚR

Závěrem je nutno podotknout, že metody pro identifikaci modelů ve spojité časové oblasti jsou numericky náročnější v porovnání s metodami pracujícími v diskrétní časové oblasti [LJU2]. Tato práce ale ukazuje, že přístup k identifikaci modelů ve spojité časové oblasti (tedy té reálné, ve které většina systémů pracuje) je výhodný především tam, kde máme nějaké informace o podstatě takového procesu (děj je popsateelný fyzikálními zákonitostmi) a hledáme pouze skupinu neznámých parametrů specifických pro daný konkrétní problém (tzv. "grey box").

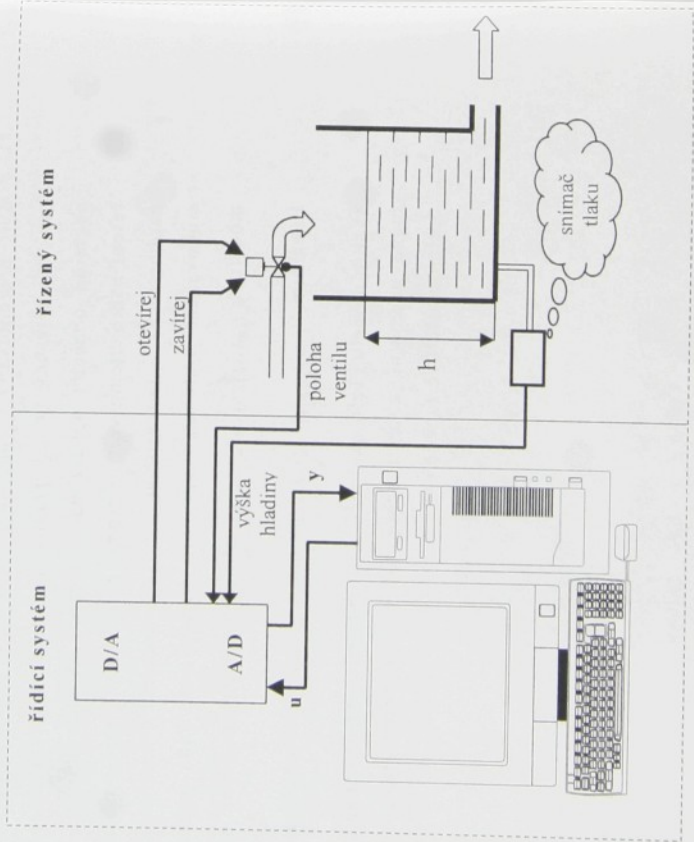
Metody identifikace modelů lineárních v parametrech tak, jak byly prezentovány v kapitolách 2.1 až 2.4, jsou vhodné především pro inicializační odhad parametrů, který je možné použít pro některou obecnou optimalizační metodu. A to z důvodu, že při řešení nahrazujeme vnitřní proměnné modelu (který je nelineární v parametrech) korespondujícími hodnotami změřenými na studovaném systému. Např. pro soustavu 1. řádu nahrazujeme model: $y_m(t, \mathbf{p}) = b \int_0^t u(\tau) d\tau - a \int_0^t y_m(\tau, \mathbf{p}) d\tau$ (nelineární v parametrech) modelem

lineárním: $y_m(t, \mathbf{p}) = b \int_0^t u(\tau) d\tau - a \int_0^t y(\tau) d\tau$, $\mathbf{p} = (a, b)$, y značí změřené hodnoty, y_m

výstup modelu. Tato okolnost, v případě silně zašuměného systému, může podstatně zhoršit schopnost těchto metod nalézt globální optimum účelové funkce.

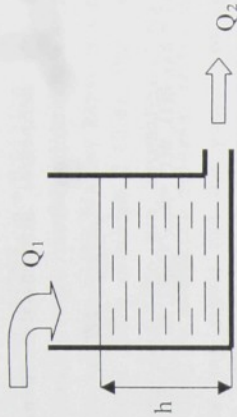
2.10 IDENTIFIKACE VÝŠKY HLADINY

Mějme hydrodynamický systém s principiálním schématem zobrazeném na následujícím obrázku (obrázek 2.10-1). Jedná se o počítačem řízenou výšku hladiny v nádrži. Styk počítače a našeho systému je realizován pomocí převodníkové karty. Výstupy z převodníkové karty jsou připojeny k regulačnímu servoventilu, který ovládáme logickými výstupy (otvírej, zavírej, stůj). Poloha ventilu je odměřována a vedena zpět přes převodníkovou kartu do počítače. Výška hladiny je snímána snímačem tlaku a tento signál je taktéž připojen na analogový vstup převodníkové karty.



obrázek 2.10-1: Výška hladiny - schéma systému

Pokusme se pomocí matematicko-fyzikální analýzy výše popsaného systému odvodit diferenciální rovnice vyjadřující dynamickou závislost výšky hladiny h kapaliny ve válcové nádobě s volným odtokem na přítékajícím množství kapaliny Q_1 (viz obrázek 2.10-2).



obrázek 2.10-2: Principiální schéma modelu výšky hladiny

Definujme:

- h ... výška hladiny [m]
- Q_1 ... přítokové objemové množství kapaliny za jednotku času [$m^3 s^{-1}$]
- Q_2 ... odtokové objemové množství kapaliny za jednotku času [$m^3 s^{-1}$]
- S_1 ... průřez nádrže [m^2]
- S_2 ... průřez výtokového otvoru [m^2]
- v_2 ... výtoková rychlost [$m s^{-1}$]
- α ... součinitel průtoku [1]
- k_{AD} ... zesílení AD převodníku [1]
- k_{DA} ... zesílení DA převodníku [1]

Při řešení předpokládáme, že se jedná o ideální kapalinu s nulovou viskozitou a $S_1 \gg S_2$. Na základě zákona zachování energie se potenciální energie ($E_p = mgh$) kapaliny uvnitř nádrže přemění na kinetickou ($E_k = \frac{1}{2}mv^2$) ve výtokovém otvoru:

$$\begin{aligned} \frac{1}{2}mv^2 &= mgh, \\ v_2 &= \sqrt{2gh}, \end{aligned} \quad (2-48)$$

výtokové množství můžeme vyjádřit takto:

$$Q_2 = \alpha S_2 v_2. \quad (2-49)$$

A jestliže napíšeme rovnici pro změnu objemu kapaliny v nádrži, pak platí:

$$\frac{dh}{dt} = S_1 \frac{dh}{dt} = Q_1 - Q_2 = Q_1 - \omega S_2 \sqrt{2g} \sqrt{h}, \tag{2-50}$$

$$h' + \frac{\omega S_2 \sqrt{2g}}{S_1} \sqrt{h} = \frac{1}{S_1} Q_1. \tag{2-51}$$

A tuto nelineární diferenciální rovnici (2-51) můžeme vyjádřit ve standardním tvaru:

$$y' = b_0 u - p \sqrt{y}, \tag{2-52}$$

kde: $y(t) = k_{AD} h(t), u(t) = k_{DA} Q_1(t), b_0 = 1/S_1, \quad p = \frac{\omega S_2 \sqrt{2g}}{S_1}.$

Z hlediska pomalé dynamiky výšky hladiny lze zanedbat dynamiku akčního členu (ventil) a snímačů. Model byl tedy uvažován ve tvaru totožném s (2-52).

Nalezené parametry integrační metodou jsou:

$$b_0 = 0.147012, \quad p = 0.012614.$$

Nalezené parametry derivační metodou jsou:

$$b_0 = 0.145316, \quad p = 0.012087.$$

Parametry nalezené těmito metodami jsou velice podobné a i odezvy těchto modelů jsou téměř identické, stejně je tomu pak i s následným vyhodnocením kvality modelu, proto jsou zde uvedeny pouze výsledky srovnání modelu a systému pro model nalezený integrační metodou.

Toto porovnání výstupu modelu získaného integrační metodou a vyšetřovaného systému výšky hladiny kapaliny v nádrži je uvedeno na následujícím grafu (obrázek 2.10-3). Pro toto srovnání byla použita stejná data jako pro optimalizaci. Výstupy modelu a systému jsou velice podobné v celé šířce výstupního signálu y a i účelová funkce má vzhledem k průměrné hodnotě signálu a délce měření malou hodnotu. Na dalším obrázku (obrázek 2.10-4) je zobrazeno srovnání reakcí modelu a systému výměniku na jiný signál než byl použit pro optimalizaci. I toto srovnání dopadlo velmi dobře a není patrný znatelný posun k horšímu, i když, jak se dalo předpokládat, kvadratická účelová funkce má o něco vyšší hodnotu. Nalezený model pro použité rozsahy vstupních a výstupních veličin dobře reprodukuje chování studovaného systému.

Nakonec ještě vyzkoušíme nalézt parametry modelu stejné struktury viz.(2-52) iterační metodou, která byla inicializována parametry získanými pomocí derivační metody.

Nalezené parametry iterační metodou:

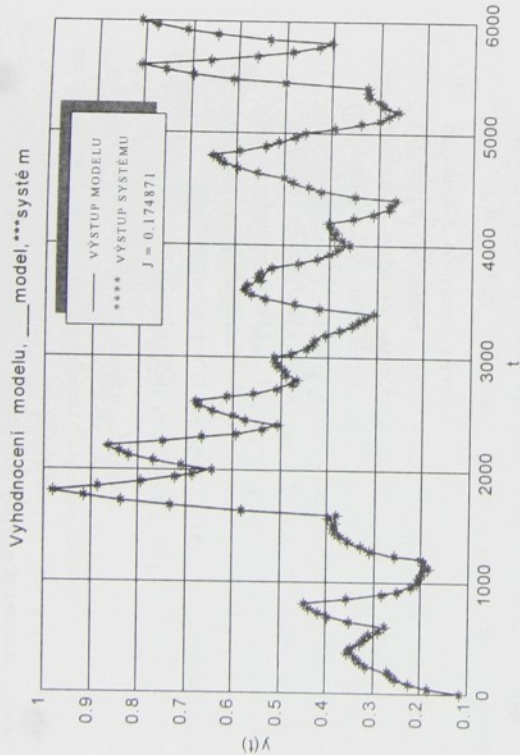
$$b_0 = 0.147293, \quad p = 0.012114.$$

Porovnání výstupu modelu získaného iterační metodou a vyšetřovaného systému výšky hladiny kapaliny v nádrži je uvedeno na následujícím grafu (obrázek 2.10-8). Pro toto srovnání byla použita stejná data jako pro optimalizaci. Na dalším obrázku (obrázek 2.10-9) je zobrazeno srovnání reakcí modelu a systému výměniku na jiný signál než byl použit pro optimalizaci. Tato srovnání dopadla velice podobně jako o modelu nalezeného integrační metodou, takže i pro tento model platí vše, co bylo řečeno v předchozím odstavci.

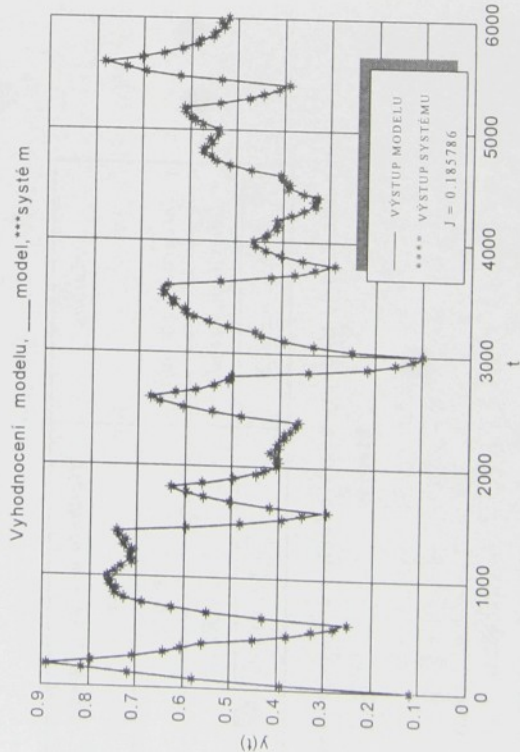
Porovnejme hodnoty kvadratické účelové funkce pro jednotlivé metody a tak zhodnotíme jak si vedly (tabulka 2.10-1). Z této tabulky vidíme, že nejlépe si vedla iterační metoda a nejhůře integrační, i když jsou mezi nimi jen nepatrné rozdíly.

Metoda	Hodnota účelové funkce $J(p)$
Integrační metoda	0.184958
Derivační metoda	0.174872
Iterační metoda	0.171594

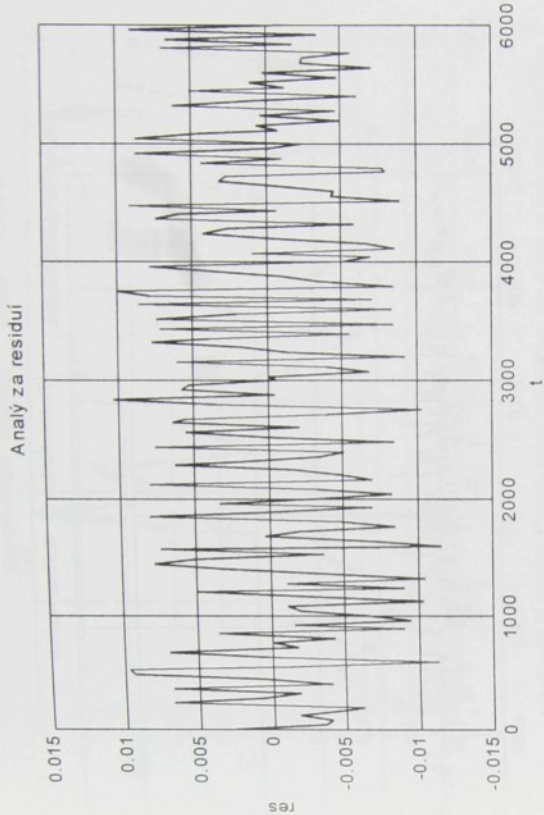
tabulka 2.10-1: Hodnoty účelové funkce pro systém výšky hladiny



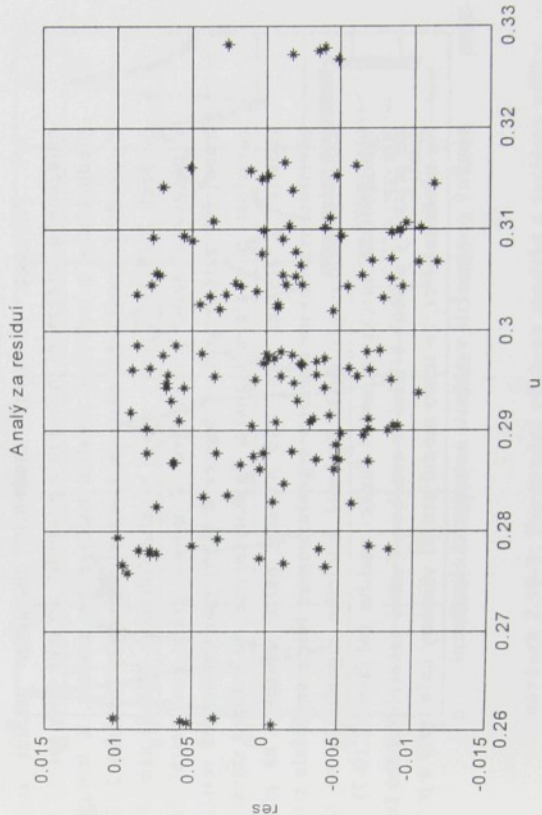
obrázek 2.10-3: Porovnání výstupu modelu a systému výšky hladiny na datech použitých pro optimalizaci



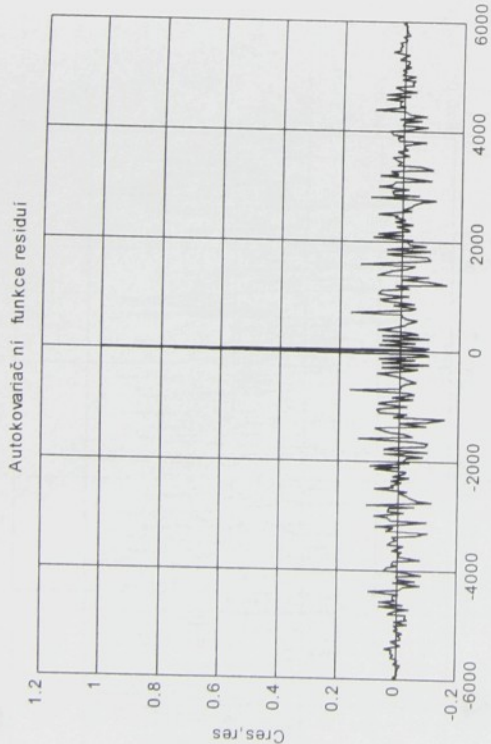
obrázek 2.10-4: Porovnání výstupu modelu a systému výšky hladiny na jiných datech než byla použita pro optimalizaci



obrázek 2.10-5: Časová posloupnost residuí modelu výšky hladiny

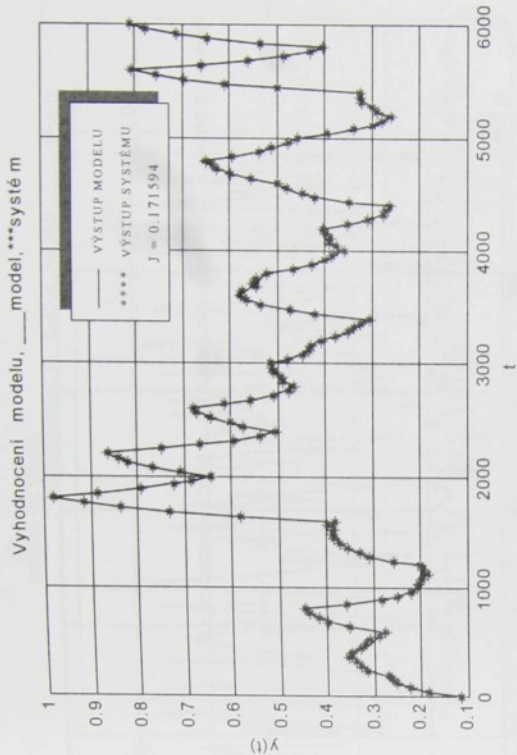


obrázek 2.10-6: Graf residuí v závislosti na budícím signálu

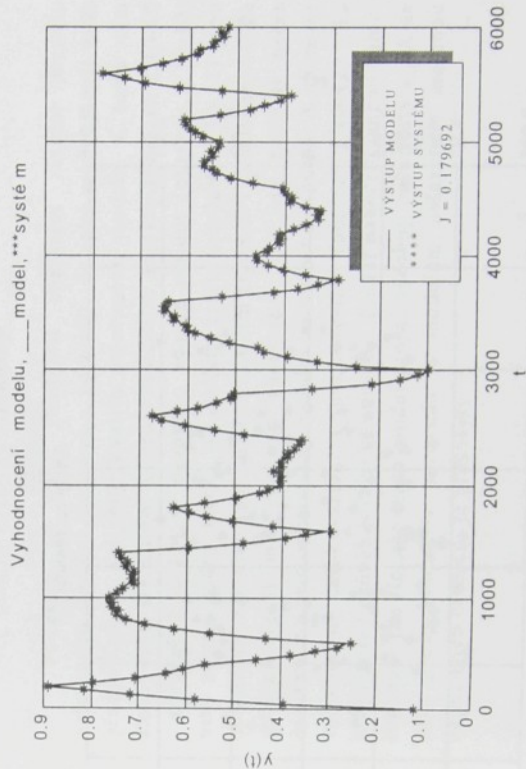


obrázek 2.10-7: Autokovarianční funkce residuí systému výšky hladiny

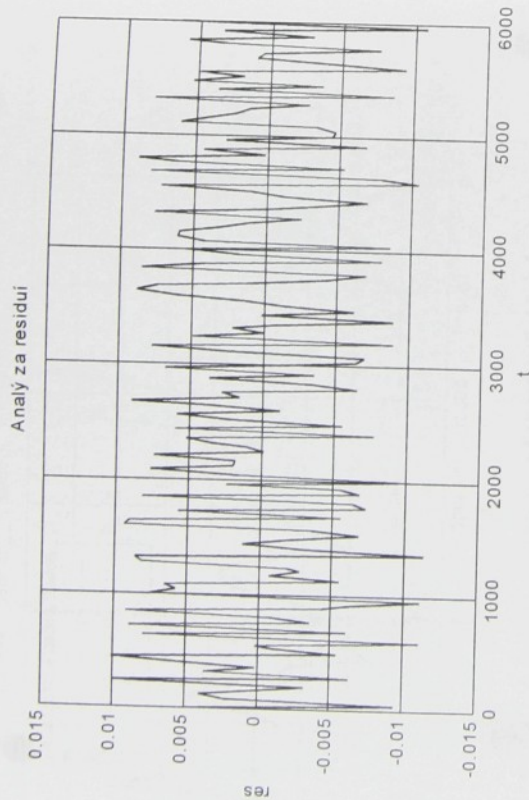
Pokusme se nyní zhodnotit charakter residuí modelu nalezeného integrační metodou a posoudit, do jaké míry splňují naše předpoklady a ovlivňují kvalitu nalezeného modelu. K tomu nám slouží další standardní nástroje residuální analýzy (obrázek 2.10-5 - obrázek 2.10-7). Časový vývoj residuí (obrázek 2.10-5) nám slouží k posouzení jejich charakteru. Z tohoto grafu není patrný žádný trend. Na dalším grafu (obrázek 2.10-6) jsou zobrazena residua v závislosti na řídicím signálu. Poznamenejme, že vstupní signál měl skokový charakter. Ani zde našťásti nelze usuzovat na přímou závislost residuí na budícím signálu; jen se lehce snižovala střední hodnota residuí při vyšší hodnotě budícího signálu. A na závěr vyhodnoňme autokovarianční funkci residuí (obrázek 2.10-7). Ta má příznivý průběh blízký se autokov. funkci nekorelovaného šumu. Na závěr lze říci, že tento model prošel všechny naše testy a je tedy vhodný pro použití v oblasti automatického řízení.



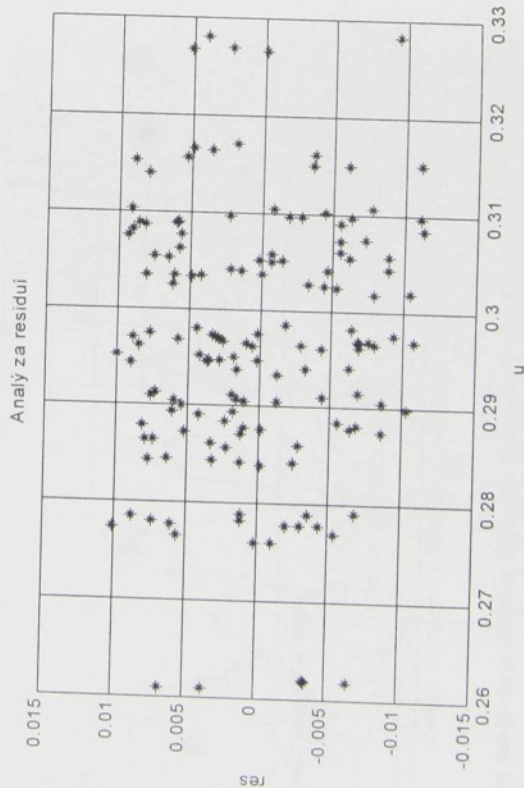
obrázek 2.10-8: Porovnání výstupu modelu a systému výšky hladiny na datech použitých pro optimalizaci



obrázek 2.10-9: Porovnání výstupu modelu a systému výšky hladiny na jiných datech než byla použita pro optimalizaci



obrázek 2.10-10: Časová posloupnost residuí modelu výšky hladiny



obrázek 2.10-11: Graf residuí v závislosti na budícím signálu



obrázek 2.10-12: Autokovarianční funkce residuí systému výšky hladiny

Pokusme se nyní zhodnotit charakter residuí modelu nalezeného iterací metodou a posoudit, do jaké míry splňují naše předpoklady a ovlivňují kvalitu nalezeného modelu. K tomu nám slouží další standardní nástroje residuální analýzy (obrázek 2.10-10 - obrázek 2.10-12). Z časového vývoje residuí (obrázek 2.10-10) není patrný žádný trend. Na dalším grafu (obrázek 2.10-11) jsou zobrazena residua v závislosti na řídícím signálu. Poznamenejme, že vstupní signál měl skokový charakter. Ani zde nastěti nelze usuzovat na přímou závislost residuí na budícím signálu. A na závěr vyhodnoťme autokovarianční funkci residuí (obrázek 2.10-12). Ta má příznivý průběh blíží se autokov. funkci nekorelovaného šumu. Na závěr lze říci, že tento model prošel všechny naše testy a jak se dalo předpokládat, v porovnání s modelem nalezeném metodou nejmenších čtverců, je přesnější.

3 Z á v ě r

Úkolem této práce bylo nalézt a vyzkoušet postupy identifikace deterministických lineárních a nelineárních parametrizovaných modelů ve spojitě časové oblasti. Je zde uvedeno odvození dvou metod (integrační a derivační) pracujících s modely lineárními v parametrech. Pro tyto dvě metody lze použít metodu nejmenších čtverců, což značně zrychluje celý výpočet. Nevýhodou těchto metod je, že při řešení nahrazujeme vnitřní proměnné modelu, který je nelineární v parametrech, korespondujícími hodnotami ze změřených dat, což může v některých případech (např. pro špatně zvolený model, či systém ovlivněný poruchou) vést k nesmyslnému výsledku. Pro tyto dvě metody je odvozeno rozšíření na soustavy s nenulovými počátečními podmínkami. Zatímco u integrační metody nenulové počáteční podmínky zvyšují počet hledaných parametrů v závislosti na řádu modelu, hledání parametrů derivační metodou tímto není nijak ovlivněno. Tyto dvě metody byly podrobeny analýze citlivosti kvality nalezeného modelu v závislosti na působící poruše a řádu modelu. Zatímco šum o vyšších frekvencích nemá tak výrazný vliv na nalezený model, šum o nízkých frekvencích již při menších úrovních působí značně destruktivně na výsledný model. Částečným řešením je rozšíření modelu o parametry trendu. Negativní vliv šumu je dále umocňován řádem identifikované soustavy (respektive modelu). Obecně si lépe vedla integrační metoda v porovnání s derivační a to platí i pro velikost vzorkovací periody, protože derivační metoda pro stejnou přesnost potřebuje menší vzorkovací periodu.

Další metodou, která se dostala do středu našeho zájmu, je metoda kvazilinearizace. Tato metoda již patří do třídy iteračních metod a pomocí náhrady nelineární funkce dvěma prvními členy Taylorova rozvoje je schopna hledat parametry nelineárních modelů. Značným omezením této metody je nutná znalost prvních parciálních derivací dle vektoru hledaných parametrů. S výhodou byla vyzkoušena kombinace integrační metody a kvazilinearizace. Toto spojení snižuje počet iterovaných parametrů pro většinu případů hledání modelů pro použití v oblasti automatického řízení.

Poslední prověřovaný postup završující skupinu metod ve spojitě časové oblasti nám umožňuje hledat parametry volně parametrizovatelných struktur dle volně definovaného kritéria.

Tento postup využívá univerzálních iteračních metod a byl úspěšně implementován v softwarovém prostředí MATLAB-SIMULINK. Značnou úsporu času výpočtu lze ve většině případů dosáhnout kombinací této metody a integrační (respektive derivační) metody, kdy integrační metoda je použita na hledání parametrů lineárního subsystému, ostatní parametry jsou počítány iteračním algoritmem.

V závěru této disertační práce byla demonstrována použitelnost těchto metod na několika skutečných systémech. Nalezené modely byly následně podrobeny kvalitativní analýze a jednotlivé metody a přístupy byly porovnány mezi sebou. Z výsledků je zřejmé, že metody identifikace ve spojitě časové oblasti mají značný potenciál pro nalezení kvalitních modelů především tam, kde něco víme o podstatě identifikovaného procesu a pouze hledáme neznámé parametry. Tyto metody si ale určitě nekladou za cíl konkurovat metodám v diskrétní oblasti (v důsledku jejich odlišnosti to ani není možné), ale spíše je mají doplňovat a v některých případech i třeba alternovat. Dalším doporučením vyplývajícím z této práce je doplňování metod používajících metodu nejmenších čtverců (ty lze použít pro počáteční odhad parametrů iterační metody) některou z iteračních metod. Tímto modifikovaným přístupem se celý postup identifikace stává daleko robustnějším vůči různým vnějším negativním vlivům.

SEZNAM LITERATURY

- [BEL] Bellman, Richard Ernest, Roth, R.: Quasilinearization and the identification problem, Singapore, World Scientific, 1983.
- [BOS] Bosch, P. J., Klaauw, A. C.: Modeling, identification, and simulation of dynamical systems, CRC Press, 1994.
- [FON] Forsman, K., Ljung, L.: On the Deadzone in System Identification, technical report, Dept. of EE. Linköping University, 1991.
- [FOR] Forsell, U., Ljung, L.: Identification for Control: Some Results on Optimal Experiment Design, technical report, Dept. of EE. Linköping University, 1998.
- [FOR1] Forsell, U.: Identification for control: unification of some existing schemes, technical report, Dept. of EE. Linköping University, 1997.
- [FOR2] Forsell, U., Lindskog, P.: Combining Semi-physical and Neural Network Modeling: An Example of Its Usefulness, Proceedings of the 11th IFAC Symposium on System Identification, pp 795-798, 1997.
- [GUN] Gunnarsson, S.: Combining Tracking and Regularization in Recursive Least Squares Identification, technical report, Dept. of EE. Linköping University, 1996.
- [GUN1] Gunnarsson, S., Hjalmarsson, H.: Some Aspects of Iterative Identification and Control Design Schemes, technical report, Dept. of EE. Linköping University, 1994.
- [HAN1] Hanuš Bořivoj - Alexik, M. - Bila, J. - Kroc, L. - Krupka, Z.: Teorie automatického řízení II, Liberec, VŠST Liberec, 1985.
- [HAN2] Hanuš Bořivoj - Balátě, J. - Švare, I. - Zíkeš, F.: Teorie automatického řízení I, 1. a 2. část, Liberec, VŠST Liberec, 1982.
- [HAN3] Hanuš Bořivoj - Balda, M.: Základy technické kybernetiky I, Liberec, VŠST Liberec, 1989.
- [HAN4] Hanuš Bořivoj - Balda, M.: Základy technické kybernetiky II. část, Brno, VUT Brno, 1981.
- [HEI] Heij, C.: Deterministic identification of dynamical systems, New York, Springer-Verlag, 1989.
- [HOR] Horáček, P., Fuka, J.: Systémy a modely, ČVUT, Praha, 1996.
- [CHE] Chen, Han-fu, Guo, L.: Identification and stochastic adaptive control, Boston, Birkhäuser, 1991.
- [JAJ] Janeček Josef - Modrák, O.: Základy technické kybernetiky, příklady, Liberec, VŠST Liberec 1990.
- [JOH] John, J., Horáček, P.: Identifikace a modelování, ČVUT, Praha, 1982.
- [LIN] Lindskog, P.: Methods, Algorithms and Tools for System Identification Based on Prior Knowledge, PhD. Thesis, Linköping University, 1996.
- [LIN1] Lindskog, P.: Fuzzy identification from a grey box modeling point of view, technical report, Dept. of EE. Linköping University, 1996.
- [LJU1] Ljung Lennart: System Identification Toolbox : For Use with MATLAB, Natic, Math Works, 1993.
- [LJU2] Ljung Lennart: System identification : Theory for the user, Englewood Cliffs, PTR Prentice Hall, 1987.
- [LJU3] Ljung Lennart: Identification for Control -- What is there to learn?, Learning, Control and Hybrid Systems, st. 207-221, 1998.
- [LJU4] Ljung Lennart: Developments for the System Identification Toolbox for MATLAB, technical report, Dept. of EE. Linköping University, 1999.
- [LJU5] Ljung Lennart: Model Validation and Model Error Modeling , technical report, Dept. of EE. Linköping University, 1999.
- [LJU6] Ljung Lennart, Guo, L.: Classical Model Validation for Control Design Purposes, Mathematical Modelling of Systems, Vol 3, st. 27-42, 1997.
- [LJU7] Ljung Lennart, Guo, L.: The Role of Model Validation for Assessing the Size of the Unmodeled Dynamics, IEEE Trans Automatic Control, Vol AC-42, st. 1230-1240, 1997.
- [LJU8] Ljung Lennart: Development of system identification, technical report, Dept. of EE. Linköping University, 1996.
- [LJU9] Ljung Lennart: System identification through the eyes of model validation, technical report, Dept. of EE. Linköping University, 1995.
- [LJU10] Ljung Lennart: Some Results on Identifying Linear Systems Using Frequency Domain Data, Proc. of the 32th IEEE Conference on Decision and Control, st. 3534-3538, 1993.
- [LJU11] Ljung Lennart, Glad, T., Andersson, T.: Identifiability implies robust identifiability, technical report, Dept. of EE. Linköping University, 1993.
- [LJU12] Ljung Lennart, Sjöberg, J.: A system identification perspective on neural nets, technical report, Dept. of EE. Linköping University, 1992.
- [LJU13] Ljung Lennart, Sjöberg, J., McKelvey, T.: On the use of regularization in system identification, technical report, Dept. of EE. Linköping University, 1992.
- [MAN] Maňas Miroslav: Optimalizační metody, Praha, SNTL, 1979.
- [MAT] Matko, D., Zupančič, B., Karba, R.: Simulation and Modelling of Continuous Systems : A Case Study Approach, Englewood Cliffs: Prentice-Hall, 1992.
- [McK] McKelvey, T.: Identification of State-Space Models from Time and Frequency Data, PhD. Thesis, Linköping University, 1995.
- [MOD1] Modrák, O.: Teorie automatického řízení, Liberec, 1992.
- [MOD2] Modrák, O.: Dynamic Systems and Feedback Control : An Introduction, Liberec, 1992.
- [MOD3] Modrák, O.: Theory of Control, Liberec, 1992.
- [NIU] Niu, S., Ljung, L., Björk, A.: Decomposition Methods for Solving Least-Squares Parameter Estimation , IEEE Trans. Signal Processing, Vol 44, st. 2847, 1996.

[NOR] Norton, J. P.: An introduction to identification, New York, Academic Press, 1985.

[OLE] Olehla, M.: Identifikace technologických soustav : Studijní podklady pro obor ASŘ ve strojírenství, TU Liberec, Liberec, 1997.

[PRE] Press William H.: Numerical recipes in C : The art of scientific computing, Cambridge, Cambridge University, 1992.

[SAH] Saha, Dines Chandra: Identification of continuous dynamical systems : the Poisson Moment Functional (PMF) approach, Springer-Verlag, 1983.

[SJO] Sjöberg, J.: Non-Linear System Identification with Neural Networks, PhD. Thesis, Linköping University, 1995.

[SJO1] Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Deylon, B., Glorennee, P., Hjalmarsen, H., Juditsky, A.: Nonlinear black-box Modeling in System Identification: a Unified Overview, technical report, Dept. of EE, Linköping University, 1995.

[SJO2] Sjöberg, J.: Some Examples of Identification with Neural Networks, technical report, Dept. of EE, Linköping University, 1994.

[SLO] Slotine, J., Weiping L.: Applied Nonlinear Control, Englewood Cliffs : Prentice-Hall, 1991.

[SOD] Söderström Torsten – Stoica, P.: System Identification, Hemel Hempstead: Prentice Hall, 1989.

[SOU] Soukup Jan: Identifikace soustav, Praha, SNTL, 1990.

[STR] Strejc, V.: Teorie lineární regulace, ČVUT, Praha, 1970.

[STR1] Strejc, V.: Teorie automatického řízení, ČVUT, Praha, 1982.

[UNB] Unbehauen, Heinz : Identification of continuous systems, New York, Elsevier Science, 1987.

[WAL] Walter Eric – Pronzato, L.: Identification of Parametric models from experimental data, Berlin, Springer, 1997.

[WIL] Willems, J. C.: From data to model, New York, Springer-Verlag, 1989.

[ZIT] Zitek Pavel: Simulace dynamických systémů, Praha, SNTL, 1990.

SEZNAM PRACÍ AUTORA SOUVISEJÍCÍ S DANOU PROBLEMATIKOU

[1] Kamler R.: Identifikace, diplomová práce, Technická Univerzita v Liberci, Liberec 1999

[2] Kamler R.: Identification, theses, University of Cincinnati, Cincinnati, OH, 1998

[3] Kamler R.: Identification, Workshop on Electronics, Control, Measurement and Signals 99, Liberec, 1999

[4] Kamler R.: Metody identifikace spojitých modelů technologických soustav, disertační práce, Technická Univerzita Liberec, Liberec 1999