

**České vysoké učení technické v Praze,
Fakulta strojní**

**Czech Technical University in Prague,
Faculty of Mechanical Engineering**

Doc.Ing.Milan Hofreiter,CSc.

**Pravděpodobnostní identifikace modelu technologického procesu pro
syntézu řízení**

**Probability Identification of Technological Process Model for Control
Synthesis**

Summary

System identification deals with the building dynamic models. Models describe relationships between measured signals. This paper concerns the selection of a proper model of the technological process appropriate for control purposes. Since the considered technological process is stochastic, probability theory is used for this purpose. In the first part of the paper the Bayesian approach to statistics (based on subjective interpretation of probability) is briefly described. The paper further concerns the selection of a proper model structure and parameter estimation within a given model structure. This is done on the basis of the observation of measured inputs and outputs of the technological process. How to apply the Bayesian approach to the identification of the technological process described by means of the linear model with unknown parameters and with the normal distribution of a stochastic component (“noise”) is presented at the end of this paper.

Souhrn

Identifikace systémů se zabývá vytvářením dynamických modelů. Modely popisují vazby mezi měřenými signály. Přednáška je věnována výběru modelu technologického procesu vhodného pro návrh řízení. Jelikož uvažovaný technologický proces je stochastický, je využita teorie pravděpodobnosti pro tento záměr. V první části přednášky je stručně popsán bayesovský přístup ke statistice (založený na subjektivním pojetí pravděpodobnosti). Další část přednášky se zabývá výběrem vhodné struktury modelu a odhadem parametrů v rámci dané struktury na základě dat změřených na technologickém procesu. Jak použít bayesovský přístup při identifikaci technologického procesu popsaného pomocí lineárního modelu s neznámými parametry a s normálním rozdělením stochastické složky („šumem“) je uvedeno na konci přednášky.

Klíčová slova: identifikace systému, model, pravděpodobnost, bayesovská statistika, technologický proces

Keywords: system identification, model, probability, Bayesian statistics, technological process

Obsah

1. Úvod	6
2. Užitá symbolika	9
3. Pravděpodobnostní popis technologického procesu	11
4. Identifikace struktury modelu stochastické soustavy a odhad jeho parametrů	14
5. Regresní model	20
5.1. Obecný regresní model	20
5.2. Lineární regresní model s gaussovskou stochastickou složkou	21
6. Závěr	27
7. Literatura	28

Equation Chapter 1 Section 1

1. Úvod

Každý z nás musí během dne učinit celou řadu rozhodnutí, která následně určují naše chování. Přitom naše rozhodnutí závisí na individuálním posouzení svých možností, situace, ve které se nalzáme, a požadovaném cíli. Jedná se o rozhodnutí typu co, kdy a jak udělat. Ve své mysli si vytváříme vědomě či nevědomě modely situací za účelem odhadu budoucích důsledků našich rozhodnutí. Výběr určité varianty závisí na požadovaném cíli a zvoleném kritériu optimality, které vědomě či nevědomě volíme a jež odráží subjektivní vyhodnocení budoucího vývoje pro naše různá rozhodnutí.

Při odhadu budoucích důsledků našich rozhodnutí vycházíme z dosavadních zkušeností. Informace, které získáváme o okolním prostředí nám pomáhají k lepšímu pochopení dějů okolního světa a k přesnějšimu odhadu budoucího vývoje. Poznávací proces zaměřený na ztotožnění představ a poznatků subjektu (pozorovatele) se zkoumaným objektem se nazývá identifikace.

Stejný postup je využíván i při návrhu řízení technologických procesů. Chceme-li technologický proces řídit musíme být schopni odhadnout jeho chování pro uvažovanou strategii řízení. Pomocí identifikace technologického procesu můžeme nalézt matematický model procesu. Matematické vyjádření popisující vztahy mezi systémovými proměnnými, např. pomocí diferenciálních nebo diferenčních rovnic, umožňuje předpovídat chování technologického procesu pro různé strategie řízení a tím i umožňuje na základě zvoleného kritéria optimality odhadnout optimální řízení.

Pro popis soustavy tj. určení matematického modelu technologického procesu lze pouze výjimečně nalézt dostatečně jednoduchý a spolehlivý model soustavy pomocí matematicko-fyzikální analýzy aplikací základních fyzikálních zákonů, neboť touto cestou získaný model bývá většinou příliš složitý a nelze ho proto použít přímo pro syntézu algoritmů řízení a je nutno ho zjednodušit. Dosud neexistuje dostatečně obecná a spolehlivá metoda aproximace, která by zaručovala požadovanou shodu chování původního a zjednodušeného modelu v uzavřeném zpětnovazebním obvodu s regulátorem, který má být teprve určen. Připočteme-li k tomu neurčitosti ve fyzikálním popisu procesu (nepřesnosti v určení empirických koeficientů, soustředění parametrů atd.) a uvážíme-li existenci náhodných vlivů s neznámými statistickými charakteristikami, které se u většiny průmyslových procesů výrazně projevují, dostaneme realistický obraz o možnostech tohoto postupu pro identifikaci technologických procesů k získání matematického modelu pro syntézu řízení.

Na rozdíl od matematicko-fyzikální analýzy praktické zkušenosti s experimentální identifikací ukazují, že se často vystačí s relativně jednoduchými modely, jestliže se jejich parametry určují přímo v pracovních podmínkách, ve kterých má být model využíván. Model procesu má představovat takové zjednodušení chování procesu, které zachovává jeho podstatné vlastnosti vzhledem použitému regulátoru a k cílům řízení. Je chybou si představovat, že existuje jediný nejlepší model procesu. Při seřizování PID regulátoru v praxi například často postačuje model, který charakterizuje zesílení s dominantní časovou konstantou. Pro řízení soustavy při velkém rozsahu pracovního bodu bychom měli znát i její nelineární charakteristiku. Zatímco pro návrh regulátoru deterministického procesu stačí deterministický model soustavy, při řízení stochastických procesů musí model vystihovat jak soustavu tak i poruchu (stochastickou složku). Vidíme, že na model jsou kladeny velmi různé požadavky. Na druhé straně naše výpočetní možnosti ovlivňují strukturu i složitost modelu a pro syntézu algoritmu řízení většinou potřebujeme znát lineární model procesu, protože složitější modely (nelineární) obecně využít neumíme.

Při matematickém modelování technologických procesů zatížených náhodnými vlivy a neurčitostmi je vhodnější popsat závislost mezi měřenými vstupy a výstupy stochastickou transformací než deterministickou závislostí, neboť deterministické metody identifikace neumožňují kvantitativně ohodnotit kvalitu odhadu parametru resp. struktury modelu a je vhodnější je aplikovat v případech, kdy vliv náhodných poruch a neurčitostí je zanedbatelný, což pro případ modelování obecného technologického procesu nelze předpokládat.

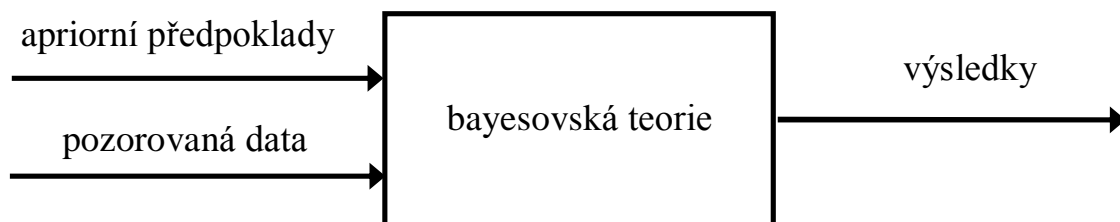
Z výše uvedeného důvodu je dále popsán pravděpodobnostní (Bayesův) přístup k identifikaci, kdy se chápou neznámé veličiny jako náhodné veličiny s daným apriorním rozložením pravděpodobnosti na množině možných jejich hodnot. Identifikace pak spočívá v určování aposteriorních rozložení pravděpodobností těchto veličin podmíněných pozorováními provedenými na soustavě. Tento přístup

- umožňuje získat popis vnějšího chování soustavy v každém časovém okamžiku
- nevyžaduje stacionaritu identifikovaného procesu
- neznámé parametry neodhaduje, ale počítá aposteriorní rozložení pravděpodobnosti těchto parametrů
- umožňuje vložit apriorní informaci
- soustřeďuje pozornost uživatele na modelování jeho specifického problému a nikoliv na výběr statistické metody

Uvedenému pravděpodobnostnímu přístupu k identifikaci je dána přednost před statistickým přístupem, neboť statistický přístup k identifikaci má některé omezení (viz.[10], [11]) např.

- většina závěrů platí pouze asymptoticky
- výsledky identifikace nejsou korektně využitelné pro syntézu regulátorů

Moderní axiomatická teorie pravděpodobnosti je exaktní nástroj, který umožňuje pracovat s matematickými modely a číselnými daty, což je pro ní vstupem. To znamená, že teoretické výsledky jsou pouze logickým důsledkem těchto vstupů. Jelikož návrhář identifikační části vnáší do ní subjektivní prvky, jako je např. předpoklad struktury, počátečních podmínek apod., musí i výsledky identifikace, kterými jsou aposteriorní rozložení pravděpodobnosti sledovaných veličin, odrážet tyto subjektivní prvky. Takto chápaná pravděpodobnost bývá označována jako subjektivní pravděpodobnost [9],[11],[13]. Z předchozího je zřejmé, že toto pojetí nezastírá, ale naopak využívá přítomnosti subjektivních prvků, je tedy vlastně objektivnější než interpretace jiné (např. chápání pravděpodobnosti jen jako relativní četnost tohoto jevu). Přitom každý teoretický výsledek bayesovské teorie je pouze logickým důsledkem apriorních předpokladů a pozorovaných dat, viz obr.1.1.



Obr.1.1

Bayesovský přístup k identifikaci náhodných a neúplně známých dynamických systémů je v České republice dlouhodobě rozvíjen v Ústavu teorie informace a automatizace Akademie věd České republiky. Zde dosažené teoretické i praktické výsledky s mezinárodním ohlasem byly prezentovány četnými články, konferenčními příspěvky, výzkumnými zprávami, semináři a přednáškami. K publikacím, které podávají ucelený přehled obdržených výsledků z této oblasti např. patří [13],[9].

Equation Chapter (Next) Section 1

2. Užitá symbolika

Uvedeme pouze nejdůležitější dohody platné v celém výkladu.

Vzhledem k snadnějšímu vyjadřování a jelikož to nemůže vest k nedorozumění, nebudeme rozlišovat označení náhodné proměnné nebo události, její realizace a příslušné proměnné v pravděpodobnostním rozdělení. Kvůli vytvoření jednotného tvaru obecných vztahů budeme používat $p(x)$, jak pro rozdělení pravděpodobnosti, jestliže x je diskrétní náhodná proměnná, tak i pro hustotu pravděpodobnosti, jestliže x je spojitá náhodná proměnná. Z téhož důvodu je třeba integrály v obecných vztazích nahradit konečným součtem, jestliže jde o integrály vztahující se k náhodné proměnné diskrétního typu, jež může nabýt jedinou hodnotu z konečného počtu možných hodnot.

$p(\cdot|\cdot)$ značí podmíněnou hustotu (rozdělení) pravděpodobnosti; argumenty dosazené na místa rezervované tečkami, budou udávat o kterou hustotu (rozdělení) pravděpodobnosti se konkrétně jedná. První argument indikuje náhodnou proměnnou (událost), které se hustota (rozdělení) pravděpodobnosti týká a argument za svislou čarou je vyhrazen pro podmínku, tj. pro vyznačení těch náhodných veličin (událostí), které jsou v uvažované hustotě (rozdělení) pravděpodobnosti fixní (jejich realizace jsou známé)

$p(x, y)$ sdružená hustota (rozdělení) pravděpodobnosti náhodných veličin x, y

$E[\cdot|\cdot]$ podmíněná střední hodnota

$\varphi = \{x_1, x_2, \dots, x_n\}$ φ je množina tvořená prvky x_1, x_2, \dots, x_n

∞ rovnost až na normalizující faktor

$\{x: V(x)\}$ značí množinu všech x , které mají vlastnost $V(x)$

t diskrétní čas

x_t veličina x v diskrétním čase t

x_i^j časová posloupnost veličiny x od diskrétního času i do diskrétního času $j \geq i$, tj. $x_i^j = x_i, x_{i+1}, \dots, x_j$; pro případ $i=1$ místo x_1^j budeme psát x^j

A^T transpozice matice A

$|A|$ determinant matice A

$tr(A)$ stopa matice A

$A > 0$ matice A je pozitivně definitní

$A \geq 0$ matice A je pozitivně semidefinitní

$\int_{\varphi_x} p(x) dx$ integrace na množině φ_x , kde φ_x značí množinu všech možných reálných hodnot x

Při odvozování potřebných vztahů budeme často využívat níže uvedené označení následujících 2 pravidel z teorie pravděpodobnosti [3],[9],[13]:

Pravidlo marginalizace (integrace, eliminace). Marginální podmíněná hustota pravděpodobnosti $p(a|c)$ může být získána z podmíněné hustoty pravděpodobnosti $p(a,b|c)$ integrací přes φ_b tj. množinu všech možných hodnot náhodné veličiny b a tedy platí

$$p(a|c) = \int_{\varphi_b} p(a,b|c) db \quad (2.1)$$

Pravidlo násobení. Sdružená hustota pravděpodobnosti $p(a,b|c)$ může být vyjádřena jako součin hustoty pravděpodobnosti $p(a|b,c)$ a hustoty pravděpodobnosti $p(b|c)$, nebo-li

$$p(a,b|c) = p(a|b,c) \cdot p(b|c) \quad (2.2)$$

Pro identifikaci technologického procesu budeme využívat bayesovský (subjektivně pravděpodobnostní) přístup, který pravděpodobnost chápe jako subjektivní míru jednotkové statistikovy důvěry rozdělené na množině možných hodnot, kterých náhodná proměnná může nabýt. V bayesovském pohledu přitom každá náhodná proměnná může nabýt pouze jednu hodnotu. Jestliže zjistíme hodnotu náhodné proměnné, pak tato proměnná se změní v číslo.

Zakladatelem uvedeného bayesovské statistiky je reverend Thomas Bayes, který je i autor tzv. **Bayesova vzorce**

$$p(a|b) = \frac{p(b|a) \cdot p(a)}{p(b)} = \frac{p(b|a) \cdot p(a)}{\int_{\varphi_a} p(b|a) \cdot p(a) da}, \quad (2.3)$$

jež snadno odvodíme pomocí vztahů (2.1) a (2.2)

Další méně častá značení zavedeme až na místech, kde je budeme potřebovat.

Equation Chapter (Next) Section 1

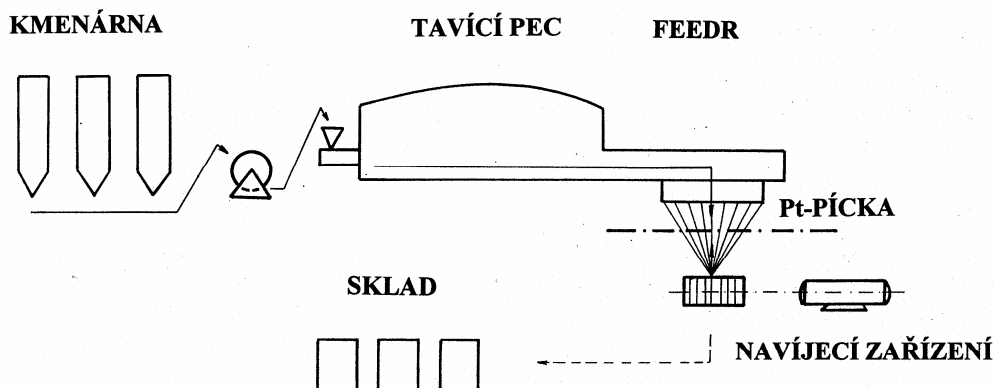
3. Pravděpodobnostní popis technologického procesu

Subjektivně pravděpodobnostní metodiku uvedenou v předchozí části na použijeme v následující části.

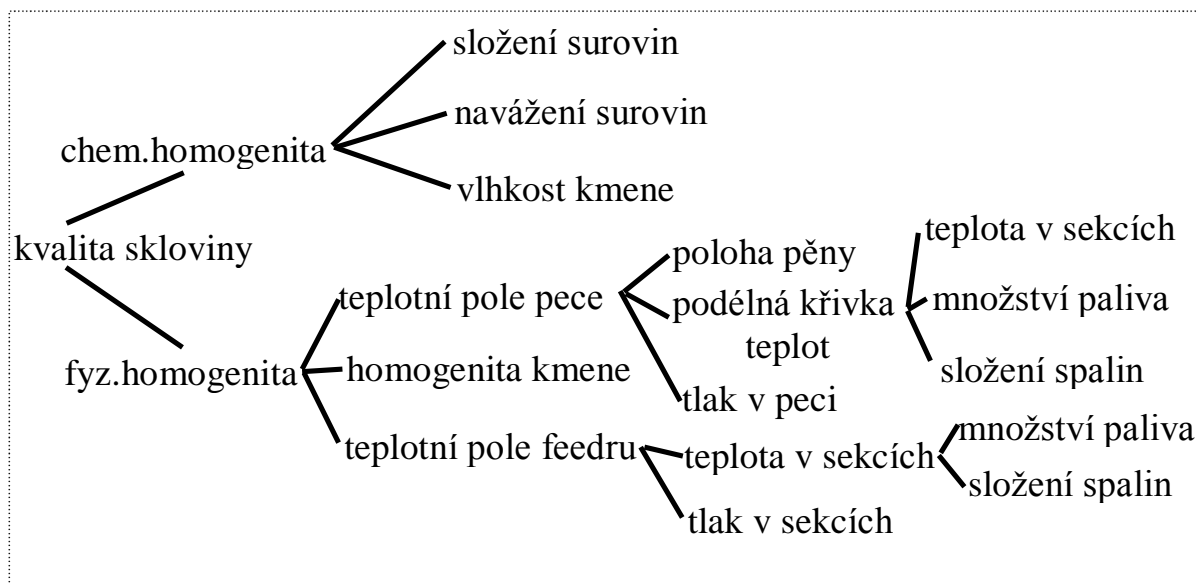
Předpokládejme, že máme za úkol řídit technologický proces, kdy úkolem automatizovaného řízení je udržovat určitý soubor regulovaných veličin na předepsaných hodnotách. Tím se zajišťuje stabilizace určitého pracovního režimu výrobního zařízení, který je z nějakého důvodu optimální a na který bylo výrobní zařízení navrženo. Je přirozenou snahou provozovatele udržovat kontinuální výrobní proces v takovém režimu co nejlépe, tj. udržovat vybraný soubor veličin v tolerancích daných normami kvality výrobku a technicko-ekonomickými požadavky provozu. Úkolem identifikace daného technologického procesu je stanovit matematický model vhodný pro syntézu řízení.

Pro názornost si můžeme představit například sklářský tavící agregát určený pro výrobu nekonečných skleněných vláken (viz. obr.3.1), kde probíhá současně velké množství interakcí a naší snahou je získat co nejvíce kvalitních informací o průběhu procesu výroby. Avšak pro jejich získání existuje prakticky pouze omezená množina parametrů sklářského tavícího agregátu, které lze zjistit měřením. Jde především o teploty, tlaky, výšku hladiny skloviny a chemické složení spalin. Přitom rozhodující vliv na jakost skloviny má nastavení optimálního teplotního profilu sklářského tavícího agregátu. Tento základní parametr je ovlivňován především jakostí topných médií, tepelnými a tlakovými podmínkami, průběhem spalování a výškou hladiny (viz. obr. 3.2).

V běžné praxi technolog předepíše, jaké mají být základní parametry sklářského agregátu v statickém režimu provozu a úkolem řízení je udržet proces v okolí předepsaného pracovního bodu.



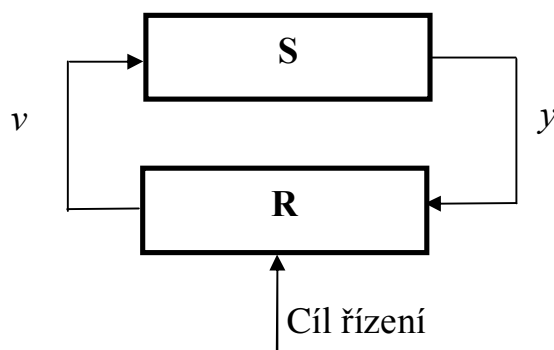
Obr.3.1 Jednostupňová výroba nekonečných skleněných vláken



Obr.3.2 Základní veličiny ovlivňující kvalitu skloviny

Pro účely obecného popisu technologického procesu proto uvažujme, že stochastická soustava S , jež je předmětem našeho studia, je řízena číslicovým regulátorem R (viz obr.3.3), který má možnost na ní působit prostřednictvím časové posloupnosti číselného μ_v -rozměrného vektoru v tj. $\{v_1, v_2, v_3, \dots\}$. Na soustavě S lze dále měřením získat časovou posloupnost μ_y - rozměrné číselné veličiny y tj. $\{y_1, y_2, y_3, \dots\}$, jež budeme nazývat výstupem procesu. Příslušné číslico-analogové převodníky a akční členy zahrneme do řízené soustavy S .

V dalším budeme předpokládat, že číslicový počítač může k výpočtu akčního zásahu v_t použít pouze číselné vektory y_{t-1}, v_{t-1} a starší, kde t značí diskrétní čas.



Obr.3.3 Stochastická soustava S řízená číslicovým regulátorem R

Každý reálný proces přesně vzato má konečné trvání. Prvnímu vzorku, který můžeme na procesu pozorovat přisudíme časový index $t = 1$ a poslednímu, který nás ještě zajímá index $t = t_K$.

Zavedme označení pro pozorované vstupy a výstupy procesu vztahující se k diskrétnímu času t vztahem (3.1)

$$D_t = \{y_t, v_t\} \quad (3.1)$$

Jestliže známe vstupní-výstupní data až do času t_0 včetně (viz Pozn.1), pak pro popis řízeného procesu pro účely návrhu číslicového řízení bychom pro libovolnou realizovatelnou strategii chtěli znát hodnoty vstupů a výstupů $\{D_{t_0+1}, D_{t_0+2}, \dots, D_{t_K}\}$, kterou dle zavedeného označení zapíšeme $D_{t_0+1}^{t_K}$. Je však zřejmé, že tento požadavek je nerealistický, neboť žádný matematický model reálného procesu nemůže dát takovou apriorní znalost s absolutní přesností a je rozumné jej zmírnit na nalezení takového modelu procesu, který umožňuje určit hustotu pravděpodobnosti $p(D_{t_0+1}^{t_K} | D^{t_0})$ pro jakoukoliv přípustnou strategii řízení.

Pozn. 1:

Předpoklad znalosti D^{t_0} není na újmě obecnosti, neboť t_0 může být rovno 0, a tedy D^{t_0} je prázdná množina.

Použitím pravidla násobení pro hustoty pravděpodobnosti můžeme hustotu pravděpodobnosti $p(D_{t_0+1}^{t_K} | D^{t_0})$ upravit na tvar

$$p(D_{t_0+1}^{t_K} | D^{t_0}) = \prod_{t=t_0+1}^{t_K} p(D_t | D^{t-1}) \quad (3.2)$$

Další úpravou využitím pravidla násobení pro hustoty pravděpodobnosti obdržíme

$$p(D_t | D^{t-1}) = p(y_t, v_t | D^{t-1}) = p(y_t | v_t, D^{t-1}) \cdot p(v_t | D^{t-1}) \quad (3.3)$$

a tedy

$$p(D_{t_0+1}^{t_K} | D^{t_0}) = \prod_{t=t_0+1}^{t_K} p(y_t | v_t, D^{t-1}) \cdot p(v_t | D^{t-1}) \quad (3.4)$$

Hustoty pravděpodobnosti, které vystupují na pravé straně vztahu (3.4) mají následující interpretaci. Hustota $p(v_t | D^{t-1})$ vyjadřuje popis strategie řízení, tj. transformaci obecně stochastickou, kterou je vstup v_t určen na základě známé minulé historie procesu. Zbývající hustota pravděpodobnosti $p(y_t | v_t, D^{t-1})$ vyjadřuje závislost výstupu y_t na známé minulé historii procesu. Množina podmíněných hustot

$$p(y_t | v_t, D^{t-1}) \quad t = t_0 + 1, t_0 + 2, \dots, t_K \quad (3.5)$$

je tedy úplným pravděpodobnostním popisem řízeného procesu z hlediska vnějšího pozorovatele.

Dále bude prováděna identifikace modelu charakterizující vlastnosti řízené soustavy tj. (3.5).

Equation Chapter (Next) Section 1

4. Identifikace struktury modelu stochastické soustavy a odhad jeho parametrů

Shrňme-li úvahy uvedené ve 3.kapitole, je zřejmé, že úkolem identifikace je určení množiny podmíněných hustot pravděpodobnosti (3.5). Pro určení těchto podmíněných hustot pravděpodobnosti je však třeba znát strukturu a parametry modelu procesu. Při identifikaci proto rozlišíme dva kroky:

- 1.výběr struktury modelu stochastické soustavy
- 2.odhad parametrů v rámci dané struktury modelu stochastické soustavy

Při hledání struktury modelu stochastické soustavy tj. v našem případě technologického procesu, se setkáváme výjimečně s tím, že lze určit strukturu modelu jednoznačně a je proto třeba nalézt postup, který by nám pomohl vybrat vhodnou strukturu modelu identifikovaného procesu.

Modelem systému M budeme rozumět jakýkoliv matematický popis soustavy, který definuje prostřednictvím konečně rozměrné množiny parametrů množinu podmíněných hustot

$$p(y_t | D^{t-1}, v_t, M) \quad \text{pro } t = t_0 + 1, t_0 + 2, \dots, t_K \quad (4.1)$$

Jelikož M se vyskytuje v podmíněné části této hustoty pravděpodobnosti, znamená to, že různé modely definují obecně různé hustoty pravděpodobnosti (4.1). Je zřejmé, že model M je specifikován strukturou a parametry modelu, kde struktura i parametry jsou zatím uvažovány jako časově invariantní.

Označme φ_M množinu všech možných modelů, které připadají v úvahu pro popis stochastické soustavy. Pak zřejmě

$$M \in \varphi_M \quad (4.2)$$

Z hlediska návrháře, který je v pozici vnějšího pozorovatele, není třeba pro predikci výstupu systému rozlišovat modely, jež definují pro $t = t_0 + 1, \dots, t_K$ stejnou množinu hustot (rozdělení) pravděpodobnosti (4.1).

Nechť \bar{M} je model, který je matematickým obrazem identifikované soustavy ve smyslu stochastické transformace mezi vstupy a výstupy. Označíme-li ${}_i H$ hypotézu, že \bar{M} má určitou strukturu a pomocí ${}_i K$ neznámé parametry tohoto modelu, potom modely této struktury s parametry z množiny možných hodnot parametrů ${}_i K$ definují třídu modelů ${}_i \varphi \in \varphi_M$. Dále předpokládejme, že množina φ_M může být rozdělena do r modelových tříd tj.

$$\varphi_M = \bigcup_{i=1}^r {}_i \varphi \quad (4.3)$$

a tedy množina všech uvažovaných hypotéz je

$$\{ {}_i H ; {}_i H \equiv \{ \bar{M} \in {}_i \varphi \} \}, i = 1, 2, \dots, r \quad (4.4)$$

V dalším budeme předpokládat, že informace řídicího systému o neznámých parametrech a struktuře modelu procesu není bohatší než informace vnějšího pozorovatele. Tyto podmínky podle [11],[13] nazveme přirozenými podmínkami řízení a lze je vyjádřit následujícím předpokladem

$$p(v_t, {}_i K, {}_i H | D^{t-1}) = p(v_t | D^{t-1}) \cdot p({}_i K, {}_i H | D^{t-1}) \quad t = t_0 + 1, t_0 + 2, \dots, t_K \quad (4.5)$$

nebo-li v_t a ${}_i H, {}_i K$ jsou podmíněně nezávislé za podmínky, že je známa minulá historie procesu obsažená v datech D^{t-1} .

Ze vztahu (4.5) je zřejmé, že ho lze též vyjádřit pomocí

$$p(v_t | {}_i K, {}_i H, D^{t-1}) = p(v_t | D^{t-1}) \quad t = t_0 + 1, t_0 + 2, \dots, t_K \quad (4.6)$$

Z pravidla násobení dostaneme

$$p({}_i H, {}_i K | v_t, D^{t-1}) \cdot p(v_t | D^{t-1}) = p(v_t | {}_i H, {}_i K, D^{t-1}) \cdot p({}_i H, {}_i K | D^{t-1}) \quad (4.7)$$

vzhledem k (4.6) pak vyplývá alternativní vztah

$$p({}_i H, {}_i K | v_t, D^{t-1}) = p({}_i H, {}_i K | D^{t-1}) \quad t = t_0 + 1, t_0 + 2, \dots, t_K \quad (4.8)$$

což při slovním vyjádření znamená, že nezískáme informaci o struktuře resp. parametrech modelu procesu pozorováním pouze vstupních veličin.

Přirozené podmínky řízení jsou splněny i když vstup v je generován v otevřené smyčce, kdy nezávisí ani na vlastnostech identifikovaného systému ani na výstupu y , ale pouze na své minulé historii.

Jak bylo již řečeno vlastní identifikace stochastické soustavy spočívá v určení množiny podmíněných hustot (3.5). V našem obecném případě však neznáme ani parametry ani strukturu modelu popisující stochastickou soustavu, ale pouze předpokládáme, že máme k dispozici množinu hypotéz ${}_i H, i = 1, 2, \dots, r$ o struktuře modelu. Pokud neznáme ani konečný počet parametrů ${}_i K$ modelu, potom takto neúplně určený model definuje pouze hustoty pravděpodobnosti

$$p(y_t | v_t, D^{t-1}, {}_i K, {}_i H) \quad t = t_0 + 1, t_0 + 2, \dots, t_K \quad (4.9)$$

což znamená, že tato hustota pravděpodobnosti náhodné veličiny y_t je zadána za podmínky, že bude známo, která hypotéza předpokládá ekvivalentní strukturu modelu s modelem \bar{M} a současně budou i známy též hodnoty příslušných parametrů.

Pro určení množiny hustot pravděpodobnosti (3.5) použitím přirozených podmínek řízení, pravidla marginalizace a pravidla násobení získáme vztah

$$p(y_t | v_t, D^{t-1}) = \int_{\Phi_M} p(y_t | v_t, D^{t-1}, M) \cdot p(M | D^{t-1}) dM \quad (4.10)$$

kde

$$M \equiv \{ {}_i K, {}_i H \} \quad (4.11)$$

Jelikož hustotu pravděpodobnosti $p(y_t | v_t, D^{t-1}, M)$ známe viz. (4.9), zbývá určit $p(M | D^{t-1})$. Využitím pravidla násobení platí

$$p(M | D^{t-1}) = \frac{p(D_{t_0+1}^{t-1} | D^{t_0}, M) \cdot p(M | D^{t_0})}{p(D_{t_0+1}^{t-1} | D^{t_0})} \propto \quad (4.12)$$

$$\propto p(M | D^{t_0}) \cdot p(D_{t_0+1}^{t-1} | D^{t_0}, M)$$

Další úpravou využitím pravidla násobení obdržíme

$$p(M | D^{t-1}) \propto p(M | D^{t_0}) \cdot \prod_{\tau=t_0+1}^{t-1} p(y_\tau, v_\tau | D^{\tau-1}, M) \quad (4.13)$$

Jelikož vzhledem k (4.6) platí

$$\begin{aligned} p(y_\tau, v_\tau | D^{\tau-1}, M) &= p(y_\tau | v_\tau, D^{\tau-1}, M) \cdot p(v_\tau | D^{\tau-1}, M) = \\ &= p(y_\tau | v_\tau, D^{\tau-1}, M) \cdot p(v_\tau | D^{\tau-1}) \end{aligned} \quad (4.14)$$

pak z (4.13) a (4.14) získáme

$$p(M | D^{t-1}) \propto p(M | D^{t_0}) \cdot \prod_{\tau=t_0+1}^{t-1} p(y_\tau | v_\tau, D^{\tau-1}, M) \quad (4.15)$$

kde člen $p(v_\tau | D^{\tau-1})$ je zahrnut v multiplikativní konstantě .

Vztah (4.15), pomocí kterého můžeme určit hledanou podmíněnou hustotu pravděpodobnosti $p(M | D^{t-1})$, obsahuje na pravé straně známé hustoty pravděpodobnosti $p(y_\tau | v_\tau, D^{\tau-1}, M)$, $\tau = t_0 + 1, t_0 + 2, \dots, t - 1$ a apriorně zvolenou hustotu pravděpodobnosti $p(M | D^{t_0})$. Apriorně zvolená hustota pravděpodobnosti pak odráží naši subjektivní míru důvěry v ten či onen uvažovaný model identifikované stochastické soustavy. Výpočet $p(M | D^{t-1})$ lze pak chápat jako korekci subjektivního apriorního rozložení pravděpodobnosti tj. $p(M | D^{t_0})$ objektivními daty.

Pro posouzení jednotlivých hypotéz o struktuře modelu stochastické soustavy na základě změřených dat D^{t-1} potřebujeme určit podmíněné rozdělení pravděpodobnosti $p({}_i H | D^{t-1})$, $i = 1, 2, \dots, r$. Pro tento účel využijeme výsledný vztah (4.15), který lze vzhledem k zavedenému označení (4.11) dále upravit pomocí pravidla eliminace následujícím způsobem

$$p({}_i H | D^{t-1}) \propto \int_{\varphi_{iK}} p({}_i K, {}_i H | D^{t_0}) \cdot \prod_{\tau=t_0+1}^{t-1} p(y_\tau | v_\tau, D^{\tau-1}, {}_i K, {}_i H) d_i K \quad (4.16)$$

Protože vzhledem k pravidlu násobení platí

$$p({}_i K, {}_i H | D^{t_0}) = p({}_i K | {}_i H, D^{t_0}) \cdot p({}_i H | D^{t_0}) \quad (4.17)$$

lze dosazením tohoto vztahu do (4.16) stanovit pro $i = 1, 2, \dots, r$

$$\begin{aligned}
p({}_i H | D^{t-1}) &\propto \\
&\propto p({}_i H | D^{t_0}) \cdot \int_{\varphi_{iK}} p({}_i K | {}_i H, D^{t_0}) \cdot \prod_{\tau=t_0+1}^{t-1} p(y_\tau | v_\tau, D^{\tau-1}, {}_i K, {}_i H) d_i K \quad (4.18)
\end{aligned}$$

Zavedeme-li označení

$${}_i I(t-1, t_0) = \int_{\varphi_{iK}} p({}_i K | {}_i H, D^{t_0}) \cdot \prod_{\tau=t_0+1}^{t-1} p(y_\tau | v_\tau, D^{\tau-1}, {}_i K, {}_i H) d_i K \quad (4.19)$$

pak

$$p({}_i H | D^{t-1}) = \frac{{}_i I(t-1, t_0) \cdot p({}_i H | D^{t_0})}{\sum_{j=1}^r {}_j I(t-1, t_0) \cdot p({}_j H | D^{t_0})} \quad (4.20)$$

Výpočet $p({}_i H | D^{t-1}), i = 1, 2, \dots, r$ lze opět chápat jako korekci našeho subjektivního apriorního rozložení pravděpodobnosti (tj. $p({}_i H | D^{t_0})$ a $p({}_i K | {}_i H, D^{t_0}), i = 1, 2, \dots, r$) na prostoru možných hodnot (hypotéz) objektivními daty.

Přirozeně vyvstává otázka : Jak volit apriorní rozložení pravděpodobnosti, chceme-li „nechat mluvit data sama za sebe“?. Naštěstí není tato otázka kritická, neboť stačí volit apriorní rozložení velmi ploché (např. normální s velkou dispersí a nulovou střední hodnotou). Chceme-li však být extrémně objektivní můžeme zvolit rozložení pravděpodobnosti rovnoměrné po celém konečném, avšak jakkoliv rozlehlém prostoru možných hodnot.

V našem případě bude celkem přirozený apriorní předpoklad, že všechny hypotézy jsou stejně možné a tedy předpokládat, že

$$p({}_i H | D^{t_0}) = \frac{1}{r} \quad \text{pro } i = 1, 2, \dots, r \quad (4.21)$$

Vzhledem k (4.21) lze vztah (4.18) pro $i = 1, 2, \dots, r$ přepsat na tvar

$$\begin{aligned}
p({}_i H | D^{t-1}) &\propto \int_{\varphi_{iK}} p({}_i K | {}_i H, D^{t_0}) \cdot \prod_{\tau=t_0+1}^{t-1} p(y_\tau | v_\tau, D^{\tau-1}, {}_i K, {}_i H) d_i K = \\
&= {}_i I(t-1, t_0) \quad (4.22)
\end{aligned}$$

Snažíme-li se být extrémně „objektivní“ i při volbě apriorní hustoty pravděpodobnosti $p({}_i K | {}_i H, D^{t_0})$ zvolíme jí s rovnoměrným rozdělením pravděpodobnosti na prostoru φ_{iK} tj. na konečně rozměrném prostoru možných hodnot ${}_i K$. Pak

$$\begin{aligned}
p({}_iK | {}_iH, D^{t_0}) &= {}_i c & \text{pro } {}_iK \in \varphi_{{}_iK} \\
p({}_iK | {}_iH, D^{t_0}) &= 0 & \text{pro } {}_iK \notin \varphi_{{}_iK}
\end{aligned} \tag{4.23}$$

kde $i=1,2,\dots,r$ a pro ${}_i c$ platí

$${}_i c = \frac{1}{\int_{\varphi_{{}_iK}} d_{{}_iK}} \tag{4.24}$$

Vztah (4.20) pak vzhledem k (4.21) a (4.24) přechází na tvar

$$p({}_iH | D^{t-1}) \propto {}_i c \cdot \int_{\varphi_{{}_iK}} \prod_{\tau=t_0+1}^{t-1} p(y_\tau | v_\tau, D^{\tau-1}, {}_iK, {}_iH) d_{{}_iK} \tag{4.25}$$

pro $i=1,2,\dots,r$

Znalost rozdělení pravděpodobnosti $p({}_iH | D^{t-1})$, $i=1,2,\dots,r$ poskytuje racionální základnu při rozhodování o vhodné volbě struktury modelu stochastické soustavy. Současně ale hned poznamenejme, že pro predikci výstupu není nutné vybrat jednu strukturu modelu, ale je možné počítat se všemi uvažovanými modelovými strukturami, které jsou uvažovány jako možné, samozřejmě ale s váhami odpovídající jejich pravděpodobnosti.

Zbývá nám odhadnout parametry modelu v rámci zvolené struktury, což z hlediska uvažovaného pravděpodobnostního přístupu k identifikaci znamená, že potřebujeme určit příslušné hustoty pravděpodobnosti $p({}_iK | D^{t-1}, {}_iH)$, $i=1,2,\dots,r$ na základě znalosti minulé historie procesu obsažené v datech D^{t-1} .

Pro určení $p({}_iK | D^{t-1}, {}_iH)$ lze opět využít obecný vztah (4.13), ze kterého vzhledem k zavedenému označení (4.11), přirozeným podmínkám řízení a pravidlu násobení plyne

$$\begin{aligned}
p({}_iK, {}_iH | D^{t-1}) &= p({}_iK | {}_iH, D^{t-1}) \cdot p({}_iH | D^{t-1}) \propto \\
&\propto p({}_iK | {}_iH, D^{t_0}) \cdot p({}_iH | D^{t_0}) \cdot \prod_{\tau=t_0+1}^{t-1} p(y_\tau | v_\tau, D^{\tau-1}, {}_iK, {}_iH)
\end{aligned} \tag{4.26}$$

Vynecháním normujícího členu nezávislém na ${}_iK$ pomocí předchozího vztahu (4.26) obdržíme hledanou hustotu pravděpodobnosti $p({}_iK | D^{t-1}, {}_iH)$

$$p({}_iK | D^{t-1}, {}_iH) = \frac{p({}_iK | {}_iH, D^{t_0}) \cdot \prod_{\tau=t_0+1}^{t-1} p(y_\tau | v_\tau, D^{\tau-1}, {}_iK, {}_iH)}{{}_i I(t-1, t_0)} \tag{4.27}$$

Výpočet $p({}_iK | D^{t-1}, {}_iH)$ podle vztahu (4.27) lze opět chápat jako korekci zvoleného apriorního pravděpodobnostního rozdělení $p({}_iK | {}_iH, D^{t_0})$ na

prostoru možných hodnot objektivními daty. Chceme-li být extrémně „objektivní“ zvolíme $p({}_i K | {}_i H, D^{t_0})$ apriorně s rovnoměrným rozdělením pravděpodobnosti na prostoru $\varphi_{{}_i K}$ tj. na konečném prostoru možných hodnot ${}_i K$ a tedy platí (4.23), (4.24). Vztah (4.27) vzhledem k předchozímu pak nabývá tvar

$$p({}_i K | D^{t-1}, {}_i H) \propto \prod_{\tau=t_0+1}^{t-1} p(y_\tau | v_\tau, D^{\tau-1}, {}_i K, {}_i H) \quad (4.28)$$

kde členy nezávislé na ${}_i K$ jsou opět vynechány.

Jelikož cílem systémového modelování je určit budoucí výstup, potom prediktivní rozdělení pravděpodobnosti může být vzhledem k přirozeným podmínkám řízení (4.5) stanoveno následujícím postupem.

$$p(y_t | v_t, D^{t-1}) = \sum_{i=1}^r p(y_t | v_t, D^{t-1}, {}_i H) \cdot p({}_i H | D^{t-1}) \quad (4.29)$$

Při určení $p(y_t | v_t, D^{t-1}, {}_i H)$ vyjdeme z následující identity

$$p(y_t | v_t, D^{t-1}, {}_i H) = \int_{\varphi_{{}_i K}} p(y_t | v_t, D^{t-1}, {}_i K, {}_i H) \cdot p({}_i K | D^{t-1}, {}_i H) d_{{}_i K} \quad (4.30)$$

Po dosazení za $p({}_i K | D^{t-1}, {}_i H)$ z (4.27) obdržíme

$$p(y_t | v_t, D^{t-1}, {}_i H) = \frac{{}_i I(t, t_0)}{{}_i I(t-1, t_0)} \quad (4.31)$$

Dosazením do (4.29) za $p(y_t | v_t, D^{t-1}, {}_i H)$ a $p({}_i H | D^{t-1})$ použitím vztahů (4.31) a (4.20) získáme hledaný vztah pro predikci výstupu

$$p(y_t | v_t, D^{t-1}) = \frac{\sum_{i=1}^r {}_i I(t, t_0) \cdot p({}_i H | D^{t_0})}{\sum_{j=1}^r {}_j I(t-1, t_0) \cdot p({}_j H | D^{t_0})} \quad t = t_0 + 1, t_0 + 2, \dots, t_K \quad (4.32)$$

Vztah (4.32) odráží již zmíněnou skutečnost, že nejsme pro predikci výstupu nuceni vybrat hypotézu o struktuře modelu stochastické soustavy.

Při konkretizaci vztahů popisující identifikaci je důležitou otázkou, zda jsme schopni pro danou úlohu získat prakticky použitelné algoritmy. Bylo by totiž nesprávné se domnívat, že principiální řešení podané v předchozí části je vždy snadno realizovatelné. Mezi základní požadavky, které je nutno splnit patří požadavek omezené paměti. Jeho splnění zaručuje, že nároky na paměť počítače a na výpočetní čas na jeden identifikační krok s časem neomezeně nerostou. Je tedy zřejmé, že je potřeba při praktickém použití všechny pozorované údaje komprimovat pomocí konečně rozměrné funkce dat.

Equation Chapter (Next) Section 1

5. Regresní model

V předchozí kapitole byly odvozeny vztahy, které lze použít při identifikaci obecného technologického procesu. Jak již bylo řečeno je při bayesovském přístupu k identifikaci nutné předem vložit určité apriorní informace o struktuře a parametrech modelu.

5.1. Obecný regresní model

Jelikož naším cílem je stanovit matematický model řízeného technologického procesu zatíženého náhodnými vlivy a neurčitostmi, jde zpravidla o popis stochastické transformace mezi vstupními a výstupními signály procesu v daném pracovním režimu. Přitom se ukazuje, že tato transformace může být v rámci prakticky dosažitelné přesnosti stejně dobře popsatelná různými typy modelů.

Jak bylo uvedeno v 3.kapitole vztahem (3.5) množina podmíněných hustot

$$p(y_t | v_t, D^{t-1}) \quad t = t_0 + 1, t_0 + 2, \dots, t_K \quad (5.1)$$

je úplným pravděpodobnostním popisem řízeného procesu z hlediska vnějšího pozorovatele.

Zavedeme-li podmíněnou střední hodnotu

$$E[y_t | v_t, D^{t-1}] \quad \text{kde } t = t_0 + 1, t_0 + 2, \dots, t_K$$

pak je obecně podmíněná střední hodnota deterministickou funkcí veličin pozorovaných na procesu až do doby t . Lze tedy psát

$$E[y_t | v_t, D^{t-1}] = r_t(v_t, D^{t-1}) \quad \text{pro } t = t_0 + 1, t_0 + 2, \dots, t_K \quad (5.2)$$

kde $r_t(v_t, D^{t-1})$ je deterministická funkce veličin v_t, D^{t-1} .

Stochastickou složku soustavy (šumem) budeme rozumět stochastický proces definovaný vztahem

$$e_t = y_t - E[y_t | y^{t-1}, v^t] \quad \text{kde } t = t_0 + 1, t_0 + 2, \dots, t_K \quad (5.3)$$

Z vlastností podmíněné střední hodnoty a z definice stochastické složky soustavy (5.3) plyne, že střední hodnota šumu e_t podmíněná znalostí minulých hodnot vstupu a výstupu je nulová a dále, že šum je posloupnost nekorelovaných veličin s nulovou střední hodnotou a je nekorelovaný s y^{t-1}, v^t .

Z (5.2) a (5.3) vyplývá, že model relace vstup - výstup lze vyjádřit ve tvaru

$$y_t = r_t(y^{t-1}, v^t) + e_t \quad \text{pro } t = t_0 + 1, t_0 + 2, \dots, t_K \quad (5.4)$$

Tento model nazýváme obecným regresním modelem a funkci $r_t(v_t, D^{t-1})$ regresní funkcí. Tím jsme rozložili náhodnou transformaci minulé historie

procesu $\{v_t, D^{t-1}\}$ na nejbližší následující výstup, na deterministickou část a náhodnou složku e_t s výše uvedenými vlastnostmi. Výběr modelu procesu lze chápat jako vhodnou volbu deterministické funkce $r_t(v_t, D^{t-1})$ a vhodný způsob vyjádření těch statistických charakteristik náhodné složky e_t , které jsou pro řešení daného problému významné.

5.2. Lineární regresní model s gaussovskou stochastickou složkou

Při identifikaci řízeného technologického procesu zatíženého náhodnými vlivy a neurčitostmi, často hledáme vhodný popis stochastické transformace mezi vstupními a výstupními signály procesu v okolí daného pracovního bodu. Hledaný matematický model přitom musí dostatečně přesně vystihnout všechny důležité vlastnosti a nemá být zbytečně komplikovaný. Jeho správná volba vychází z cíle identifikace tj. v našem případě stanovit matematický model procesu pro účely řízení, kdy nejčastěji cílem řízení technologických procesů je udržet vybrané technologické parametry na požadovaných hodnotách.

Pro výše uvedené podmínky se v praxi pro popis stochastické soustavy (pro účely řízení) osvědčil lineární časově invariantní vícerozměrný regresní model s gaussovskou stochastickou složkou (dále tzv. základní model).

Základní model lze napsat ve tvaru

$$y_t = P^T \cdot z_t + e_t \quad \text{pro } t = t_0 + 1, t_0 + 2, \dots, t_K \quad (5.5)$$

kde y_t je výstup soustavy (v rozměrný sloupcový vektor) v diskrétním čase t

P je matice parametrů, jež má ρ řádků a v sloupců

z_t je ρ -rozměrný regresní vektor (sloupcový) v diskrétním čase t

e_t je v diskrétním čase t šum soustavy (v-rozměrný sloupcový vektor), který je gaussovský s konstantní kovarianční maticí R (matice s v řádky a v sloupci) a mající vlastnosti, které vyplývají s definice (5.3).

Předpokládáme, že pro regresní vektor platí

$$z_{t+1} = f(z_t, y_t, v_{t+1}) \quad \text{pro } t = t_0 + 1, t_0 + 2, \dots, t_K \quad (5.6)$$

kde f je funkce veličin z_t, y_t, v_{t+1} a platí, že z_{t_0+1} je známý vektor.

Avšak i při splnění podmínky, že lze daný technologický proces popsat základním modelem, zbývá určit řády, dopravní zpoždění, významné vstupy atd., nebo-li stanovit strukturu základního modelu. Zvolením určité struktury základního modelu jsou určeny veličiny obsažené v regresním vektoru, což znamená, že hypotéza ${}_i H, i = 1, 2, \dots, r$ o struktuře základního modelu předpokládá určitý tvar vektoru pozorování ${}_i z, i = 1, 2, \dots, r$, kde r označuje počet srovnávaných hypotéz. Jak v takovém případě postupovat bylo obecně ukázáno v kapitole 5.

V následující části je použit vyložený obecný postup pro odhad parametrů $K = \{P, \sigma^2\}$ základního modelu se známou strukturou zadaného rovnicí (5.5),

$$\text{kde } v = 1, \quad (5.7)$$

$$z_t^T = [v_t, y_{t-1}, v_{t-1}, \dots, y_{t-n}, u_{t-n}, 1], \quad (5.8)$$

tj. při daných parametrech závisí y_t jen na z_t a ne na starších datech

e_t je skalární náhodná veličina (šum soustavy) s normálním rozdělením $N(0, \sigma^2)$, tj. se střední hodnotou 0 a rozptylem σ^2 a tedy

$$p(y_t | y^{t-1}, v^t, K) = p(y_t | z_t, P, \sigma) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma^2}} \cdot \exp \left\{ -\frac{1}{2 \cdot \sigma^2} (y_t - z_t^T P)^2 \right\}. \quad (5.9)$$

Jestliže v čase $t_0 = 0$ nemáme k dispozici žádná změřená data, pouze apriorní informace o parametrech modelu, pak ze vztahu (4.27) je zřejmé, že platí

$$p(K | D^t) \propto p(K | D^0) \cdot \prod_{\tau=1}^t p(y_\tau | v_\tau, D^{\tau-1}, K), \quad (5.10)$$

kde ve vztahu byl vypuštěno rozlišení hypotézy o struktuře, neboť v tomto případě je struktura známá.

Součin modelů, vyjádřený jako funkce parametrů (zde od času 1 do času t) je věrohodnostní funkce a platí

$$\begin{aligned} \prod_{\tau=1}^t p(y_\tau | v_\tau, D^{\tau-1}, K) &= \prod_{\tau=1}^t p(y_\tau | z_\tau, P, \sigma) = \\ &= (2\pi\sigma^2)^{-\frac{t}{2}} \cdot \exp \left\{ -\frac{1}{2 \cdot \sigma^2} \sum_{\tau=1}^t (y_\tau - z_\tau^T P)^2 \right\}. \end{aligned} \quad (5.11)$$

Podle pravidel maticového počtu lze vztah (5.11) dále upravit na tvar

$$\begin{aligned} \prod_{\tau=1}^t p(y_\tau | z_\tau, P, \sigma^2) &= \\ &= (2\pi\sigma^2)^{-\frac{t}{2}} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \cdot \begin{bmatrix} -1 & P^T \end{bmatrix} \cdot \sum_{\tau=1}^t \left(\begin{bmatrix} y_\tau \\ z_\tau \end{bmatrix} \begin{bmatrix} y_\tau & z_\tau^T \end{bmatrix} \right) \cdot \begin{bmatrix} -1 \\ P \end{bmatrix} \right\}. \end{aligned} \quad (5.12)$$

Zavedeme-li z výpočetních důvodů apriorní hustotu pravděpodobnosti $p(K | D^0)$ v podobném tvaru, tj.

$$p(K | D^0) = \sigma^{-\frac{\kappa(0)}{2}} \cdot \exp \left\{ -\frac{1}{2 \cdot \sigma^2} \begin{bmatrix} -1 & P^T \end{bmatrix} V(0) \begin{bmatrix} -1 \\ P \end{bmatrix} \right\}, \quad (5.13)$$

kde $V(0)$ je pozitivně definitní symetrická matice odpovídajícího stupně a $\kappa(0)$ je skalár.

Ze vztahů (5.10), (5.12) a (5.13) po úpravě získáme

$$p(K | D^t) \propto \sigma^{-\frac{\kappa(t)}{2}} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \begin{bmatrix} -1 & P^T \end{bmatrix} V(t) \begin{bmatrix} -1 \\ P \end{bmatrix} \right\}, \quad (5.14)$$

kde

$$V(t) = V(t-1) + \begin{bmatrix} y_t \\ z_t \end{bmatrix} \begin{bmatrix} y_t & z_t^T \end{bmatrix} = V(0) + \sum_{\tau=1}^t \left(\begin{bmatrix} y_\tau \\ z_\tau \end{bmatrix} \begin{bmatrix} y_\tau & z_\tau^T \end{bmatrix} \right), \quad (5.15)$$

$$\kappa(t) = \kappa(t-1) + 1 = \kappa(0) + t \quad (5.16)$$

Rozdělíme-li matici $V(t)$ následujícím způsobem

$$V(t) = \begin{bmatrix} V_y(t) & V_{yz}(t)^T \\ V_{yz}(t) & V_z(t) \end{bmatrix}, \quad (5.17)$$

kde $V_y(t)$ je skalár, $V_{yz}(t)$ je sloupcový vektor s ρ prvky a $V_z(t)$ je čtvercová matice s ρ řádky a ρ sloupci. Pomocí tohoto rozdělení můžeme z (5.14) získat

$$p(K | D^t) \propto \sigma^{-\frac{\kappa(t)}{2}} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \begin{bmatrix} P - \hat{P}(t) \end{bmatrix}^T V_z(t) \begin{bmatrix} P - \hat{P}(t) \end{bmatrix} + \Lambda(t) \right\}, \quad (5.18)$$

kde

$$\hat{P}(t) = V_z(t)^{-1} \cdot V_{yz}(t), \quad (5.19)$$

$$\Lambda(t) = V_y(t) - V_{yz}(t)^T \cdot V_z(t)^{-1} \cdot V_{yz}(t) \quad (5.20)$$

Ze vztahu (5.18) můžeme určit bodové odhady parametrů P, σ^2 pomocí určení polohy, pro které tato hustota pravděpodobnosti nabývá svého maxima. Pro bodový odhad regresních koeficientů P platí (5.19) a pro určení bodového odhadu rozptylu šumu σ^2 lze použít derivace k určení extrému. Odhad rozptylu šumu je

$$\hat{\sigma}^2 = \frac{\Lambda(t)}{\kappa(t)} \quad (5.21)$$

Jelikož základním cílem modelu je predikce výstupu systému určíme též prediktivní hustotu pravděpodobnosti $p(y_t | v_t, D^{t-1})$. Vzhledem přirozeným podmínkám řízení (4.5), využitím vztahů (2.1), (2.2) pro uvažovaný regresní model se známou strukturou platí

$$p(y_t | v_t, D^{t-1}) = \int_{\Phi_K} p(y_t | z_t, K) \cdot p(K | D^{t-1}) dK. \quad (5.22)$$

Provedeme-li uvedenou integraci (viz např. [1],[11]) získáme prediktivní hustotu pravděpodobnosti ve tvaru

$$p(y_t | v_t, D^{t-1}) = \frac{\Gamma\left(\frac{\kappa(t) - \rho + 2}{2}\right)}{\frac{\kappa(t)+2}{2} \cdot \pi^\rho} \cdot |V_z(t)|^{\frac{1}{2}} \cdot \Lambda(t-1)^{\frac{\kappa(t)-\rho+3}{2}}, \quad (5.23)$$

kde Γ je gama funkce.

Pro bodový odhad výstupu v čase t při neznámých parametrech modelu použijeme podmíněnou střední hodnotu $E\left[y_t | v_t, D^{t-1}\right]$, pro kterou vzhledem k (2.1), (2.2), (4.8) a (5.9) platí

$$\begin{aligned} E\left[y_t | v_t, D^{t-1}\right] &= \\ &= \int_{\Phi_{y_t}} y_t p(y_t | v_t, D^{t-1}) dy_t = \int_{\Phi_{y_t}} y_t \int_{\Phi_K} p(y_t, K | v_t, D^{t-1}) dK dy_t = \\ &= \int_{\Phi_{y_t}} y_t \int_{\Phi_K} p(y_t | z_t, K) \cdot p(K | D^{t-1}) dK dy_t \end{aligned}$$

Záměnou integrace, využitím vztahu (5.9) a pomocí úprav získáme vztah

$$\begin{aligned} E\left[y_t | v_t, D^{t-1}\right] &= \int_{\Phi_K} \left\{ \int_{\Phi_{y_t}} y_t p(y_t | z_t, K) dy_t \right\} p(K | D^{t-1}) dK = \\ &= \int_{\Phi_K} z_t^T P \cdot p(K | D^{t-1}) dK = z_t^T \int_{\Phi_K} P \cdot p(P, \sigma | D^{t-1}) dP d\sigma = . \quad (5.24) \\ &= z_t^T \int_{\Phi_P} P \cdot p(P | D^{t-1}) dP = z_t^T \cdot E\left[P | D^{t-1}\right] = z_t^T \cdot \hat{P}_{t-1} \end{aligned}$$

Pro bodovou předpověď výstupu označenou \hat{y}_t tedy platí

$$\hat{y}_t = E\left[y_t | v_t, D^{t-1}\right] = z_t^T \cdot \hat{P}_{t-1} = \hat{P}_{t-1}^T \cdot z_t, \quad (5.25)$$

kde \hat{P}_{t-1} je odhad regresních koeficientů z minulého kroku.

Jestliže kromě parametrů $K = \{P, \sigma\}$ by bylo i nutné u regresního modelu odhadnout např. řády dopravní zpoždění, významné vstupy atd., pak je nutné uvažovat hypotézy o struktuře modelu a použít postup uvedený v kapitole 5.

Následující simulační příklad ukazuje jak využitím bayesovského přístupu k identifikaci lze odhadnout parametry a strukturu základního modelu.

Příklad:

Identifikace systému popsaného rovnicí

$$y_t = v_t + 0.7y_{t-1} + 0.7v_{t-1} - 0.2y_{t-2} + 0.4v_{t-2} + e_t, \quad (5.26)$$

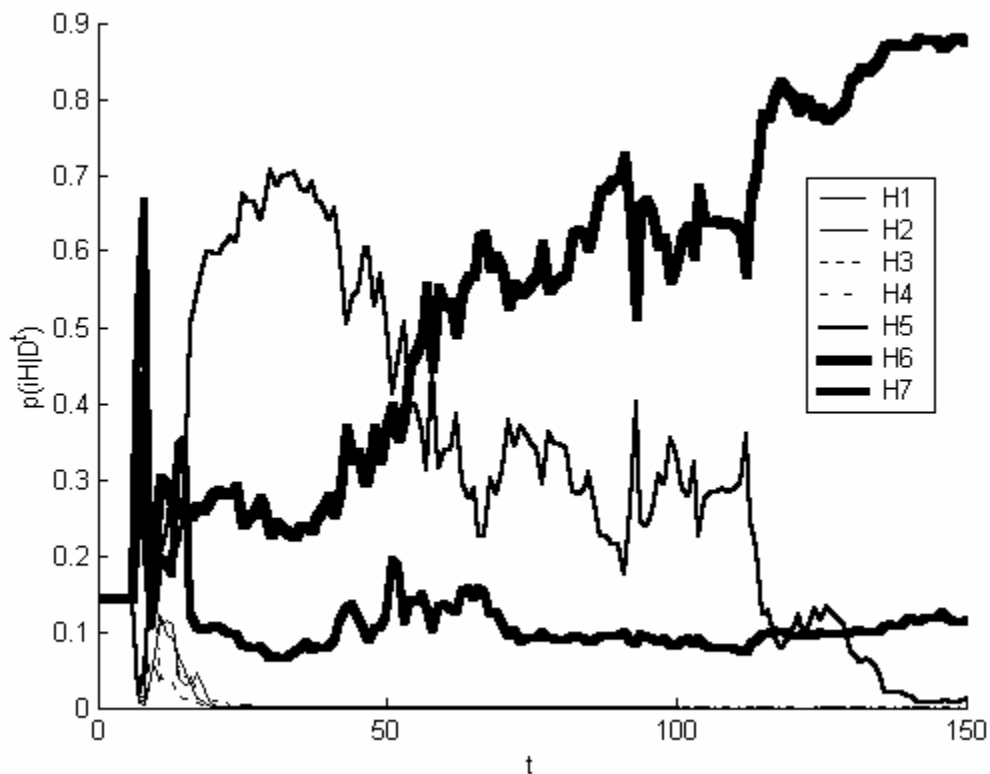
kde v i e byly generovány jako normální bílé šумы s nulovou střední hodnotou a s jednotkovým rozptylem, byla prováděna v otevřené smyčce. Úkolem bylo odhadnout na základě změřených 150 údajů o hodnotách vstupu v a výstupu y strukturu regresního modelu a v rámci struktury i parametry modelu, přičemž

byly uvažovány hypotézy ${}_iH, i=1,2,\dots,7$ o tvaru regresního vektoru z_t , kde přehled porovnávaných hypotéz je uveden v tabulce Tab. 5.1.

hypotéza	regresní vektor
${}_1H$	$z = [y_{t-1}]$
${}_2H$	$z = [y_{t-1}, y_{t-2}]$
${}_3H$	$z = [y_{t-1}, y_{t-2}, y_{t-3}]$
${}_4H$	$z = [v_t, y_{t-1}, y_{t-2}]$
${}_5H$	$z = [v_t, y_{t-1}, v_{t-1}, y_{t-2}]$
${}_6H$	$z = [v_t, y_{t-1}, v_{t-1}, y_{t-2}, v_{t-2}]$
${}_7H$	$z = [v_t, y_{t-1}, v_{t-1}, y_{t-2}, v_{t-2}, v_{t-3}]$

Tab. 5.1 Přehled porovnávaných hypotéz

Při apriorně zvoleném $\kappa(0)=10$ a $V(0)=10 \cdot I$, kde I značí jednotkovou matici, jsou průběžně určené hodnoty $p({}_iH | D^t)$, $i=1,2,\dots,7$ pro $t=1,6,\dots,150$ vykresleny na obr.5.1. Z obrázku je zřejmé, že pro $t=150$ má největší důvěru hypotéza ${}_6H$.



Obr. 5.1 Pravděpodobnosti jednotlivých hypotéz o struktuře modelu

Odhad parametrů ${}_iK = \{ {}_iP, {}_i\sigma^2 \}$ pro i -tou hypotézu o struktuře modelu je popsán hustotou pravděpodobnosti $p({}_iK | D(t), {}_iH)$, která má obdobně jako v (5.18) tvar

$$p({}_iK | D^t, {}_iH) \propto {}_i\sigma^{-\frac{\kappa(t)}{2}} \cdot \exp \left\{ -\frac{1}{2 {}_i\sigma^2} [{}_iP - {}_i\hat{P}(t)]^T {}_iV_z(t) [{}_iP - {}_i\hat{P}(t)] + {}_i\Lambda(t) \right\} \quad (5.27)$$

Z (5.27) pak můžeme případně určit bodové odhady parametrů ${}_iK = \{ {}_iP, {}_i\sigma^2 \}$ pro $i=1,2,\dots,r$ pomocí určení polohy, pro které hustoty pravděpodobnosti (5.27) nabývají svého maxima. Získané vztahy pro tyto bodové odhady jsou analogické vztahům (5.19) a (5.21) doplněné indexem i určující příslušnou strukturu modelu. V Tab. 5.2 jsou uvedeny bodové odhady parametrů ${}_iK = \{ {}_iP, {}_i\sigma^2 \}$ v diskrétním čase $t=150$ pro uvedený příklad.

$y_t = b_0v_t + a_1y_{t-1} + b_1v_{t-1} + a_2y_{t-2} + b_2v_{t-2} + a_3y_{t-3} + b_3v_{t-3} + e_t$								
	b_0	a_1	b_1	a_2	b_2	a_3	b_3	σ^2
${}_1H$	x	0,73	x	x	x	x	x	2,98
${}_2H$	x	1,07	x	-0,46	x	x	x	2,36
${}_3H$	x	1,02	x	-0,34	x	-0,11	x	2,33
${}_4H$	1,03	1,01	x	-0,40	x	x	x	1,53
${}_5H$	1,06	0,82	0,65	-0,27	x	x	x	1,39
${}_6H$	1,08	0,70	0,79	-0,25	0,40	x	x	1,23
${}_7H$	1,09	0,72	0,77	-0,23	0,36	x	-0,09	1,23

Tab. 5.2 Bodové odhady parametrů ${}_iK = \{ {}_iP, {}_i\sigma^2 \}$ v diskrétním čase $t=150$ (x znamená, že příslušná hypotéza neuvažuje odpovídající veličinu v regresním vektoru)

Equation Chapter (Next) Section 1

6. Závěr

Bylo by nesprávné se domnívat, že použitý pravděpodobnostní přístup k identifikaci technologického procesu je vždy snadno realizovatelný. Nejdále se tento přístup zatím podařilo rozpracovat pro systémy popsatelné lineárními regresními modely, které se i v praxi osvědčily (např. [4],[5],[6],[7]). Uvedený přístup je obecně použitelný i pro silně nelineární a časově variantní systémy. Jedním z perspektivních pravděpodobnostních modelů pro univerzální popis (pro účely řízení) silně nelineárního systému může být např. řízený markovský řetězec konečného řádu (viz. např. [2],[3],[14]), který lze použít i pro časově variantní systémy.

Velkou předností bayesovského odhadování je, že jde o konsistentní způsob uvažování, který umožňuje nerozporně využít teoretickou, experimentální a expertní znalost, a poskytuje odhady neznámých veličin i s informací o přesnostech těchto odhadů a to dokonce i v konečných časech pozorování.

Uvedený postup však neznamená, že je třeba zamítat intuitivní „inženýrská“ řešení, která se osvědčila, ale naopak využít uvedenou teorii pro jejich zdůvodnění a zobecnění.

7. Literatura

- [1] Anderson T.W.: An introduction to multivariate statistical analysis, John Wiley and Sons, New York, 1958 (viz též ruský překlad Gos. izd. fiziko-matem. lit., Moskva 1963).
- [2] Gao H., Kárný M.: Can Markov chains supplement adaptive AR predictors?, Preprints of the 2nd European IEEE Workshop on Computer Intensive Methods in Control and Signal Processing, pp.246-249, Prague, 1996
- [3] Hofreiter M.: Approximate Identification of Stochastic Systems, International Journal Automation Austria, Vol.5, No.2, pp. 51-63, IFAC-Beirat Österreich, 1997
- [4] Hofreiter M.: Regulace teplot ve feedru, 5.kap. závěrečné zprávy "Automatizovaný systém řízení výroby skleněného vlákna ", Výzkumná zpráva, SVÚS Hr.Králové,pobočka, Praha,1983
- [5] Hofreiter M.: Návrh strategie pro řízení spalovacího procesu", Výzkumná zpráva úkolu "ASŘTP Float II", SVÚS Hr.Králové,pobočka Praha,1988.
- [6] Hofreiter M.,Hrubý V.: Identifikace statických a dynamických lineárních systémů v reálném čase na číslicovém počítači, Výzkumná zpráva úkolu "Identifikace statistických a dynamických lineárních systémů v reálném čase na číslicovém počítači", SKLO UNION, k.ú.o.RACIS Teplice, 1984
- [7] Hofreiter M.: Experimental identification of a glass furnace, Proceedings of the Workshop 96,CTU Prague, pp.193-194,1996
- [8] Kárný M.: Bayesian estimation of model order, Problems of Control and Information Theory,no.9,pp.33-46,1980.
- [9] Nagy I.: Základy bayesovského odhadování a řízení, ČVUT, Praha, 2003
- [10] Nagy I., Nedoma P., Kárný M., L. Pavelková, Ettler P.: O bayesovském učení, AUTOMA, č.7, str. 56-60, 2002
- [11] Peterka V.: Číslicové řízení procesů s náhodnými poruchami a neurčitými charakteristikami, Doktorská disertační práce, ÚTIA ČSAV, Praha,1975.
- [12] Peterka V.,Kárný M.: Bayesian system classification, Preprints 5th IFAC Symposium on Identification and Parameter Estimation, pp.349-356, Darmstadt,1979.
- [13] Peterka V.: Bayesian system identification, Trends and Progress in System Identification, Eykhoff P.,(Ed.), pp.239-304, Pergamon Press, Oxford,1981
- [14] Sahin A.D., Sen Z.: First-order Markov chain approach to wind speed modelling, Journal of Wind Engineering and Industrial Aerodynamics 89, pp. 263-269, 2001

Doc. Ing. Milan Hofreiter, CSc.

M.Hofreiter se narodil 1.března 1952 v Praze. Po ukončení SVVŠ byl v r.1970 přijat na ČVUT v Praze, fakultu strojního inženýrství, kde studoval směr regulační a automatizační techniku. Po úspěšném obhájení diplomové práce nastoupil v r.1975 do Státního výzkumného ústavu sklářského Hradec Králové, pobočka Praha, oddělení ASŘ. V r.1978 byl ústavem vyslán na účelovou aspiranturu na ČVUT - FSI v oboru technická kybernetika, úsek automatizace ve strojírenství. V r.1981 ukončil účelovou aspiranturu v oboru technická kybernetika a obhájil disertační práci a 4.2.1982 získal na fakultě strojní ČVUT v Praze vědeckou hodnost CSc.

Po skončení aspirantury a navrácení do SVÚS jako vědecký pracovník v oddělení ASŘ řešil a vedl výzkumné úkoly SVÚS. Výzkumné úkoly, které zde převážně řešil, byly zaměřené na oblast identifikace, modelování a řízení sklářských tavících agregátů. V této oblasti dosáhl význačných výsledků a to jak v oblasti teoretické, tak i v oblasti aplikací výsledku výzkumu v průmyslové praxi (16 oponovaných výzkumných zpráv, 6 průmyslem potvrzených realizací, 4 patenty). Pro zvýšení kvalifikace absolvoval řadu odborných kursů, seminářů a školení.

Ve výzkumném ústavu, i po oddělení pobočky Praha a jejího přejmenování na Tepelná technika Praha, pracoval až do 31.5.1993, kdy na základě konkursu byl přijat jako odborný asistent na fakultu strojní ČVUT katedru automatického řízení. Svou 18-letou výzkumnou praxi využil při přednáškách a cvičeních, při vedení diplomantů resp. doktorandů a při inovaci studia na fakultě strojní ČVUT (zavedl předmět Stochastické procesy v inženýrské praxi zařazený od r. 1994 mezi povinně volitelné předměty letního semestru 4.ročníku oboru Automatické řízení a inženýrská informatika; přebudoval laboratoř Automatického řízení určenou pro všechny studenty 3.ročníku; navrhl a spolurealizoval virtuální laboratoř Automatického řízení, zavedl volitelný Seminář z automatického řízení určený pro studenty 3. a 4.ročníku; je spoluautorem přednáškových skript z automatického řízení a vedoucím autorem dvou příkladových skript pro předmět Automatické řízení). V r. 1998 byl jmenován docentem pro obor Technická kybernetika. Na fakultě strojní přednáší a cvičí především předměty "Automatické řízení" a "Stochastické procesy v inženýrské praxi". Externě dva roky zajišťoval přednášky a cvičení předmětu „Teorie automatického řízení“ na Technické fakultě ČZU v Praze. Jeho odborná aktivita je zaměřena na oblast identifikace, modelování a řízení technologických procesů. Účastní se a vede řešení grantů a výzkumných úkolů.

Je autorem resp. spoluautorem více než 62 publikací, z nichž 26 je v angličtině, 1 v ruštině a ostatní v češtině. U většiny publikací je uveden jako

jediný resp. vedoucí autor. Schopnost řešit a realizovat dosažené výsledky odráží 23 oponovaných výzkumných zpráv.

V současné době odborně spolupracuje především s ŠKODA AUTO a.s., Ústavem teorie a automatizace AV ČR, Ústavem informatiky AV ČR, státním podnikem Tepelná technika Praha a firmou Tronic Control s.r.o.

V r. 2001 byl zvolen do akademického senátu fakulty strojní ČVUT v Praze, je členem rady Ústavu přístrojové a řídicí techniky a od r.2003 vykonává funkci proděkana pro zahraniční styky.