

PROHLÁŠENÍ

Předkládám tímto k posouzení a obhajobě diplomovou práci zpracovanou na závěr studia na Fakultě aplikovaných věd Západočeské univerzity v Plzni.

Prohlašuji, že jsem diplomovou práci vypracoval samostatně a výhradně s použitím odborné literatury a pramenů, jejichž úplný seznam je její součástí.

V Plzni dne 10. 5. 2015

.....

Poděkování

Děkuji Ing. Jindřichu Duníkovi, Ph.D. za vstřícnost a poskytnutí odborné pomoci při vedení této práce.

Dále děkuji zaměstnancům společnosti Honeywell za zpracování a vyhodnocení zaslaných dat.

Anotace

Tato práce pojednává o odhadu vlastností poruch, které působí na systém. Uvažován je lineární dynamický stochastický systém popsáný stavovým modelem. Nejprve je ukázáno, jak popsat systém pomocí stavové reprezentace, jak v tomto popisu působí poruchy a jak se dají tyto poruchy reprezentovat. Dále jsou diskutovány různé metody, které rozdílnými způsoby dokážou odhadnout vlastnosti poruch působící na systém. Dále je detailně diskutována metoda založená na analýze posloupnosti chyby predikce měření pro libovolný lineární filtr a je ukázán a analyzován vliv volby lineárního zisku na kvalitu odhadu. V závěru je uvažovaná metoda rozšířena pro složitější poruchy měření ve smyslu jejich struktury a je diskutována nestrannost a konvergence této metody.

Klíčová slova: Allanova variance, Kalmanův filtr, kovariance, korelace, Kroneckerova algebra, identifikace systémů, inovační posloupnost, odhad stavu

Abstract

The thesis is devoted to the estimation of characteristics of the noises affecting a system. In particular, the linear stochastic dynamic system described by the state-space model is considered. The state-space model is introduced and techniques for the noise covariance matrices estimation are reviewed with the special emphasis on the autocovariance least-squares method. Then, the autocovariance least-squares method is analysed and a technique for the optimal selection of the linear estimator gain with respect to the purpose of the autocovariance method is proposed. The method is, subsequently, extended and analytically derived for a more general class of the systems, namely for the linear systems with time-correlated measurement noise. The proposed techniques and algorithms are illustrated in numerical illustrations.

Keywords: Allan variance, Kalman filter, covariance, correlation, Kronecker algebra, system identification, innovation sequence, state estimation

Obsah

Obsah	1
Seznam obrázků	3
Seznam tabulek	3
1. Úvod	4
2. Stavový popis dynamického systému	7
2.1. Stavový deterministický popis systému	7
2.2. Diskretizace deterministického popisu systému	9
2.3. Popis náhodné veličiny a procesu	10
2.4. Stavový stochastický popis systému	19
3. Odhad stavu stochastického systému	20
3.1. Kalmanův filtr	20
3.2. Kalmanův filtr – filtrace.....	20
3.3. Kalmanův filtr – predikce	24
4. Motivace a cíl práce	26
4.1 Předpoklady pro Kalmanův filtr	26
4.2 Odhad kovariančních matic poruch na základě chyby odhadu lineárního filtru 27	
4.3 Cíl práce.....	27
5. Autokovarianční metoda pro odhad kovariančních matic poruch LS	28
5.1 Odvození metody.....	28
5.2 Vlastnosti metody	35
6. Kovarianční matice odhadu a její horní mez.....	41
6.1 Odvození metody.....	41
6.2 Numerická ilustrace	44
7. Odhad časově korelovaných poruch.....	47

7.1	Odvození metody.....	47
7.2	Vlastnosti metody.....	51
7.3	Rozšířená metoda ACLS pro statický systém.....	52
7.4	Numerická ilustrace.....	52
8.	Softwarová implementace metody ACLS.....	58
8.1	Obecné informace.....	58
8.2	Postup a možnosti zadávání.....	59
9.	Zhodnocení.....	61
	Příloha.....	62
I.	Metoda lineárních nejmenších čtverců.....	62
II.	Metoda nelineárních nejmenších čtverců.....	63
III.	Kroneckerova algebra.....	66
	Reference.....	67

Seznam obrázků

Obrázek 1 – Hustota pravděpodobnosti rovnoměrného rozdělení	11
Obrázek 2 – Distribuční funkce rovnoměrného rozdělení	11
Obrázek 3 - Hustota pravděpodobnosti normálního rozdělení.....	12
Obrázek 4 – Distribuční funkce normálního rozdělení	12
Obrázek 5 – Histogram realizací normální veličiny.....	14
Obrázek 6 – Variance odhadu v závislosti na volbě L	36
Obrázek 7 - Variance odhadu v závislosti na volbě Q_A a R_A	37
Obrázek 8 - Variance odhadu v závislosti na volbě N a počtu kovariančních rovnic	38
Obrázek 9 - Variance odhadů Q a R a zisk L v závislosti na volbě Q_A a R_A	45
Obrázek 10 - Hodnota kritéria J v závislosti na volbě Q_A a R_A , $n=1$	45
Obrázek 11 - Hodnota kritéria J v závislosti na volbě Q_A a R_A , $n=3$	46
Obrázek 12 – Odhady rozšířené metody ACLS pro různá N , dynamický systém	53
Obrázek 13 - Odhady modifikované metody ACLS, systém bez dynamiky	54
Obrázek 14 – Odmocnina Allanovy variance v závislosti na délce úseku, logaritmické osy..	55
Obrázek 15 - Postup programem.....	59
Obrázek 16 - Ukázka výsledku odhadu z programu	60
Obrázek 17 – Znázorněný jeden krok metody nelineárních nejmenších čtverců.....	64

Seznam tabulek

Tabulka 1 – Výsledky metody Allanovy variance a rozšířené metody ACLS	57
---	----

1. Úvod

Moderní metody zpracování signálu, detekce poruch či automatického řízení vyžadují detailní matematický model uvažovaného systému či procesu. Kvalita modelu je klíčová pro získání kvalitního odhadu signálu, poruchy nebo řízení. Špatný popis systému či procesu, jehož výsledkem je nesprávný model, který je použitý při zpracování signálu nebo automatickém řízení, v lepších případech způsobí zhoršení kvality odhadu signálu nebo řízení. V horších případech způsobí nesprávný odhad signálu nebo v případě automatického řízení i nestabilitu řízeného systému.

Matematický model popisující systém se skládá jak z deterministické, tak i stochastické části, jak bylo diskutováno v [20] a [22]. Deterministická část popisuje vztahy či zákony, které jsou známy v dané oblasti, jež model popisuje. Tyto zákony mohou být například z oblasti fyziky, chemie, kinematiky nebo ekonomiky. Stochastická část popisuje či kvantifikuje neznalost či neurčitost působící na systém, kterou lze reprezentovat jako poruchu. Tato porucha působí jak na rovnici popisující dynamiku systému, tak i na měření systému. Porucha působící na rovnici dynamiky má mnoho opodstatnění. Je to například chyba modelu, které jsme se dopustili při vytváření matematického modelu z neznalosti popisovaného systému, či například zjednodušení výsledného modelu. Také se může jednat o poruchu, kterou nelze předpovídat či popsat, nebo o poruchu, která vzniká při generování řízení, tj. pokud chceme vygenerovat určitý akční zásah, ale skutečný akční zásah bude mít jinou hodnotu v důsledku nepřesnosti akčního členu. Problém s touto stochastickou částí lze spatřit jak v popisu či modelu poruchy, tak i ve způsobu jeho nalezení.

At' už bude popis systému deterministický nebo při zavedení neurčitosti stochastický, je důležité, jakým způsobem bude tento model reprezentován. Matematický model lze reprezentovat dvěma způsoby. Jedním je vstupně-výstupní reprezentace, kterou ale v této práci využívat nebudeme. Druhým způsobem je pak stavová reprezentace, která nám umožní díky zavedení stavové proměnné zjistit o systému mnohem více informací, tj. systém detailněji namodelovat. Také je nutno podotknout, že bylo navrženo mnoho technik automatického řízení a detekce poruch pro systém popsané stavovým modelem.

Jak již bylo řečeno výše, deterministická část systému se dá zjistit pomocí fyzikálních či matematických vztahů či zákonů. Otázkou tedy zůstává, jak zjistit vlastnosti poruch neboli stochastické části modelu. Pro popis těchto poruch se obvykle

využívá metody identifikace a statistiky. Od 70. let minulého století existuje snaha zjistit popis poruch pro stavově reprezentované popisy systému. Od té doby vzniklo mnoho metod, které se dají rozdělit do těchto kategorií: korelační metody [1]-[5] a [21], metody odhadu založené na Bayesovo přístupu či aproximace [6] a [13], metody založené na maximální věrohodnosti [7], kovarianční metody [8], metody založené na Minmax přístupu [9], podprostorové metody [10], metody založené na metodě chyby predikce [11], metody využívající pro odhad rozšířeného Kalmanova filtru [12]. Kromě přímého odhadu kovariančních matic poruch byly navrženy i metody, jež odhadují přímo zisk lineárního filtru jako například v [1], [14], [15]. Všechny výše uvedené přístupy mají ovšem mnoho předpokladů, vlastností a omezení, které jsou pojednávány například v [3], [9], [16] a [17]. Je důležité říci, že tyto metody jsou v mnoha případech numerické a tím pádem ani v mnoha případech nelze dokázat nestrannost odhadu metody. Proto v této práci bude také diskutována a rozvíjena metoda využívající chyby odhadu lineárního filtru navržená v [3]. Tato autokovarianční metoda pro odhad kovariančních matic poruch, v anglické psané literatuře označovaná autocovariance least-squares method, zkráceně ACLS, je analyticky odvozená a poskytuje na rozdíl od většiny ostatních metod nestranné odhady.

Metoda ACLS má také své vlastnosti a omezení. Jedním z významných omezení je i to, že metoda využívá lineární prediktor, u kterého je nutno zvolit zisk. Tato volba je omezená tím, že zisk musí být volen tak, aby chyba odhadu byla stabilní. Volba zisku také výrazně ovlivňuje kvalitu odhadu hledaných kovariančních matic poruch. Avšak doposud tato volba byla ponechána na uživateli. Je také důležité zdůraznit, že jak všechny přístupy v předchozím odstavci, tak i metoda ACLS byly navrženy pouze pro lineární časově invariantní systémy, na které působí poruchy mající bílý charakter jak v systému, tak i v měření. Tento předpoklad bývá ale omezující například v popisu, kdy porucha měření je časově korelovaná. Tímto případem jsou například inerciální senzory. Proto v této práci bude také ukázáno jak metodu ACLS rozšířit tak, aby porucha měření nemusela být pouze bílá, ale aby mohla být také časově korelovaná.

Tato práce bude tedy členěna do následujících kapitol. V kapitole Stavový popis dynamického systému bude ukázáno jak namodelovat stochastický systém. V kapitole Odhad stavu stochastického systému bude ukázáno jak odhadnout stav stochastického systému. V další kapitole Motivace a cíl práce budou diskutovány cíle

a motivace této práce. V kapitole Autokovarianční metoda pro odhad kovariančních matic poruch bude ukázána samotná metoda. V kapitole Kovarianční matice odhadu a její horní mez bude ukázána volba zisku lineárního filtru. V kapitole Odhad časově korelovaných poruch bude ukázáno jak metodu ACLS rozšířit pro korelované poruchy působící na měření. V předposlední kapitole Softwarová implementace metody ACLS bude představena metoda ACLS implementovaná v programu Matlab. V poslední kapitole Zhodnocení budou diskutovány výhody a vlastnosti metody ACLS.

2. Stavový popis dynamického systému

Tato kapitola se zabývá stavovým popisem dynamického systému, jeho diskretizací a také popisem vlastností poruchami působící jak na stav, tak i na měření.

2.1. Stavový deterministický popis systému

At' už chceme regulovat nebo například sledovat nějaký systém, musíme v kybernetickém světě znát popis tohoto systému. Tyto systémy mohou mít různý počet vstupů a výstupů. Mohou být lineární či nelineární, mohou být deterministické či stochastické a takto by mohly být systémy rozdělovány do mnoha kategorií. Tyto systémy mohou být popsány různými způsoby jako v [19], kde mezi důležité popisy lze zařadit matematické modely. Ty lze určit dvěma způsoby. Jedním z nich je matematicko-fyzikální modelování, kde ze znalosti systému lze vytvořit model a to například pomocí fyzikálních zákonů, za pomoci Newtonovské či Lagrangeovské mechaniky, ekonomických nebo chemických zákonů či vztahů popisující oblasti vytvářeného modelu. Pomocí těchto metod lze určit jak strukturu, tak i parametry modelu popisující libovolný systém. Druhým způsobem je využití identifikačních metody jako v [20]. V tomto případě si zvolíme strukturu modelu s neznámými parametry, u které věříme, že odpovídá popisu reálného systému nebo mu alespoň do určité míry dostačuje. Tato struktura poté poslouží v identifikační metodě k nalezení takových parametrů, aby výsledný model s těmito parametry se choval co nejvíce podobně reálnému systému. Obě tyto metody – matematicko-fyzikální modelování i identifikační přístup – se dají zkombinovat takovým způsobem, abychom pomocí první metody získali strukturu modelu, do které budou následně druhou metodou odhadnuty parametry, které se například nedají zjistit pomocí matematicko-fyzikálního modelování. Takto vytvořený matematický model je obvykle reprezentován dvěma způsoby.

Jedním z nich vstupně-výstupní reprezentace. Tato reprezentace, jak už jméno naznačuje, popisuje vztah mezi vstupem a výstupem systému. Pro příklad si uvedeme spojitý lineární deterministický systém popsáný diferenciální rovnicí n-tého řádu s jedním výstupem vstupem:

$$\sum_{i=0}^n (a_i * y^{(i)}(t)) = \sum_{i=1}^m (b_i * u^{(i)}(t)), \quad (2.1)$$

kde y je výstup systému, u je vstup systému, a parametry a a b jsou parametry modelu. Dále pak horní index u vstupu či výstupu je příslušná časová derivace a dolní index u parametru a a b ukazuje, ke které derivaci vstupu či výstup patří daný parametr.

Druhou možností reprezentace matematického modelu je stavová neboli vnitřní reprezentace. Tento popis zavádí novou proměnnou, a to stav x . Tento stav reprezentuje vnitřní vazby v modelu. Pro příklad uvedeme stavovou reprezentaci spojitého lineárního deterministického systému n -tého řádu s p výstupy a nu vstupy, který může vypadat následovně:

$$\begin{bmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_n(t) \end{bmatrix} = \begin{bmatrix} & & \\ & A_S & \\ & & \end{bmatrix} * \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} + \begin{bmatrix} & & \\ & B_S & \\ & & \end{bmatrix} * \begin{bmatrix} u_1(t) \\ \vdots \\ u_{nu}(t) \end{bmatrix} \quad (2.2)$$

$$\begin{bmatrix} y_1(t) \\ \vdots \\ y_p(t) \end{bmatrix} = \begin{bmatrix} & & \\ & C_S & \\ & & \end{bmatrix} * \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix} \quad (2.3)$$

$$\begin{bmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{bmatrix}, \quad (2.4)$$

kde rovnice (2.2) se nazývá stavová rovnice, ve které matice $A_S \in \mathbb{R}^{n \times n}$ reprezentuje matici dynamiky systému a matice $B_S \in \mathbb{R}^{n \times nu}$ prezentuje vstupní matici. Dále pak rovnice (2.3) se nazývá rovnice měření, ve které matice $C_S \in \mathbb{R}^{np \times n}$ reprezentuje výstupní matici. Rovnice (2.4) reprezentuje počáteční stav. Tato reprezentace se využívá velice často, neboť oproti vstupně-výstupní reprezentaci umožňuje detailní model systému reprezentující realitu včetně vnitřních vazeb systému. Další výhodou je například možnost využití různých metod detekce, identifikace, odhadu a řízení založených na stavových modelech.

Je také důležité říci, že mezi vstupně-výstupní a stavovou reprezentací existují určité vztahy. Pro příklad budeme chtít převést rovnici (2.1) do stavové reprezentace. Lze to provést například vhodnou volbou stavů x a převést na n diferenciálních rovnic prvního řádu, které představují stavové rovnice. Tyto rovnice lze následně seskupit do matic. Tyto matice spolu s maticemi, které popisují pomocí obecných funkcí stavových proměnných výstup systému, tvoří stavově popsany systém. Těchto převodů však existuje nekonečně mnoho. Lze tedy dojít k jiným maticím A_S , B_S , C_S , které popisují jeden a ten samý systém popsany obecnou rovnicí (2.1). Pokud budeme chtít převést systém popsany stavově v rovnicích (2.2) a (2.3) do vstupně-výstupní

prezentace, jako je například rovnice (2.1), lze tento převod udělat jednoznačně. To znamená, že různé stavové reprezentace popisující totožný systém lze převést na jednu rovnici, která by byla totožná s (2.1).

2.2. Diskretizace deterministického popisu systému

V dnešní době již vše od výpočtu řízení až k měření výstupu systému budeme provádět na diskretním počítači čili v diskretních okamžicích. Proto spojité modely, které byly ukázány v předchozí kapitole 2.1, je vhodné převést do diskretní podoby.

Pokud například rovnici (2.1) převedeme do přenosové funkce pomocí Laplaceovy transformace, vznikne nám přenos, který vypadá následovně:

$$F(s) = \frac{\mathcal{L}\{y(t)\}}{\mathcal{L}\{u(t)\}} \Big|_{n.p.p} = \frac{Y(s)}{U(s)} = \frac{\sum_{i=0}^m [b_i * s^i(t)]}{\sum_{i=0}^n [a_i * s^i(t)]} \quad (2.5)$$

kde funkce \mathcal{L} značí Laplaceovu transformaci a funkce Y a U jsou Laplaceovy obrazy výstupu a vstupu. Tento přenos lze jednoduše převést do diskretní podoby například lichoběžníkovou neboli Tustinovou aproximací, kterou provedeme tak, že za komplexní proměnnou s dosadíme:

$$s = \frac{2}{T} * \frac{z - 1}{z + 1} \quad (2.6)$$

kde z značí komplexní proměnnou dopředného posuvu v čase a T je perioda vzorkování. Výhodou této aproximace například oproti obdélníkové dopředné či zpětné aproximace je ten, že celou stabilní levou polorovinu komplexní roviny s převádí do stabilní jednotkové kružnice v rovině z s počátkem v nule. Rovnice (2.5) převedená pomocí rovnice (2.6) do diskretní podoby lze následně analogicky jako u spojitého systému převést do normálních forem stavových reprezentací.

Diskretizovat se dá také přímo spojitá stavová reprezentace (2.2) a to pomocí následujících vztahů:

$$A_D = e^{A * T} \quad (2.7)$$

$$B_D = A_S^{-1} * (A_D - I_{n \times n}) * B_S \quad (2.8)$$

Maticе C_S v rovnici (2.3) zůstává nezměněna. Vztah (2.6) lze ovšem provést pouze u systémů, u nichž matici A_S lze invertovat. Stavová rovnice (2.2) a (2.3) by

tedy vypadala, pokud A_D , B_D a C_D budeme pro zjednodušení psát jako A , B a C , následovně:

$$\begin{bmatrix} x_1(k+1) \\ \vdots \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} A \end{bmatrix} * \begin{bmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{bmatrix} + \begin{bmatrix} B \end{bmatrix} * \begin{bmatrix} u_1(k) \\ \vdots \\ u_{nu}(k) \end{bmatrix} \quad (2.9)$$

$$\begin{bmatrix} y_1(k) \\ \vdots \\ y_p(k) \end{bmatrix} = \begin{bmatrix} C \end{bmatrix} * \begin{bmatrix} x_1(k) \\ \vdots \\ x_n(k) \end{bmatrix}, \quad (2.10)$$

kde proměnná k , značí pro jednoduchost jednotlivé časové okamžiky v době vzorkování neboli měření výstupu systému.

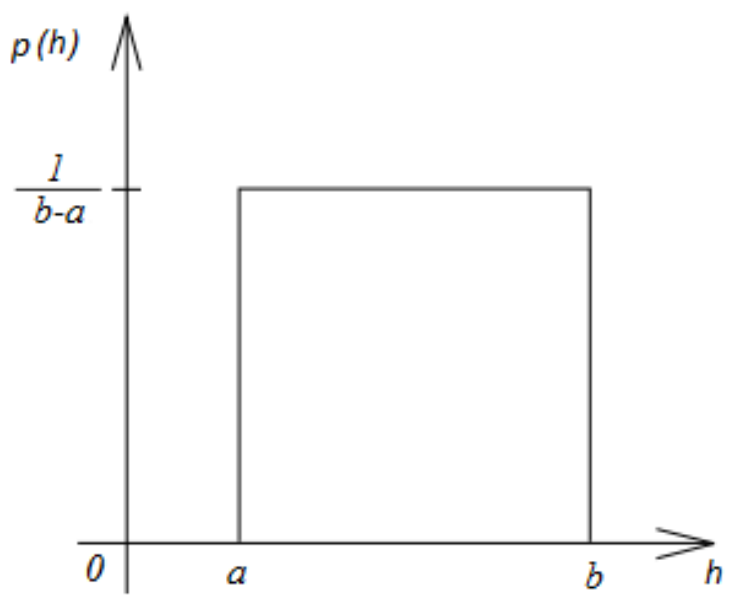
2.3. Popis náhodné veličiny a procesu

Na začátek si zavedeme náhodnou veličinu H , jež je funkcí z prostoru jevů Y do prostoru reálných čísel \mathbb{R} . Výsledek realizace této náhodné veličiny H označíme jako h . Pokud existuje míra pravděpodobnosti spojená s náhodou veličinou H tak, aby podmnožina $A \in Y$ měla dané pravděpodobnosti $P(A)$ nezáporné a aby pravděpodobnost celého prostoru $P(Y)$ byla rovná jedné, existuje distribuční funkce $F(h)$ a hustota pravděpodobnosti $p(h)$. Pro distribuční funkci platí, že je vždy rostoucí, její limita pro h blížíící se mínus nekonečnu je rovna nule a limita blížíící se nekonečnu je rovná jedné. Pro hustotu pravděpodobnosti platí, že je nezáporná, její plocha je rovná jedné a pravděpodobnost $P(a < H < b)$ je rovna určitému integrálu od a do b přes hustotu pravděpodobnosti $p(h)$. Hustota pravděpodobnosti a distribuční funkce úplně popisují náhodnou veličinu H . Mezi distribuční funkcí a hustotou pravděpodobnosti existuje také vztah, který umožňuje přecházet z jednoho popisu na druhý, a to takovýto:

$$p(h) = \frac{dF(h)}{dh}. \quad (2.11)$$

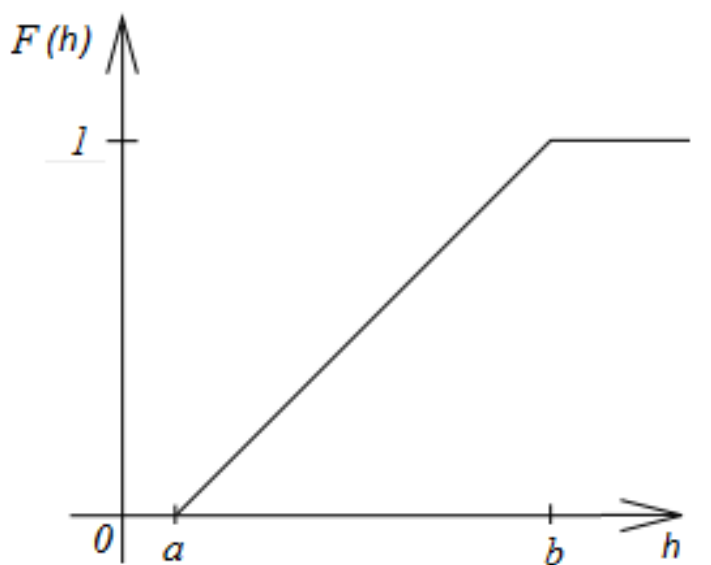
Pro příklad řekněme, že náhodná veličina H má rovnoměrné rozdělení, tj. h se vyskytuje mezi hodnotami a a b se stejnou pravděpodobností a všude jinde s pravděpodobností nula.

Hustota pravděpodobnosti této náhodné veličiny bude vypadat následovně:



Obrázek 1 – Hustota pravděpodobnosti rovnoměrného rozdělení

kde pravděpodobnost na intervalu a až b je rovna převrácenému rozdílu $b - a$ proto, aby jak už bylo definováno, byla plocha hustoty pravděpodobnosti rovna jedné. Také uvedeme distribuční funkci, která vypadá následovně:

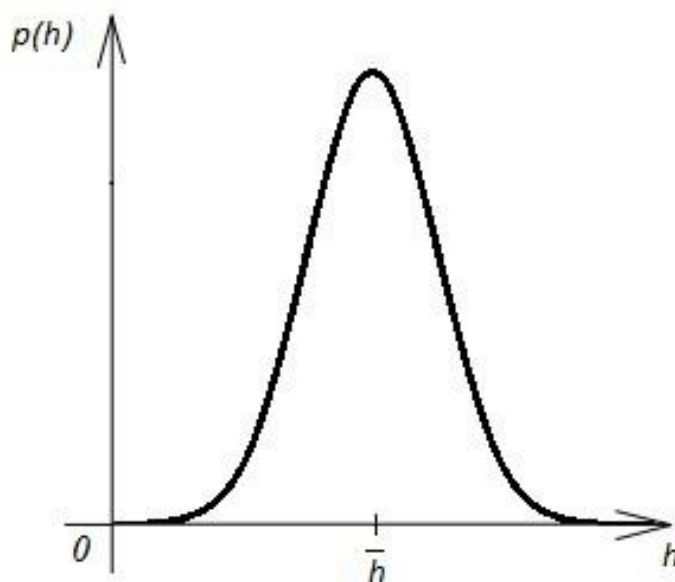


Obrázek 2 – Distribuční funkce rovnoměrného rozdělení

Pro příklad nyní řekněme, že náhodná veličina H má gaussovské rozdělení. Toto rozdělení je popsané následujícím vzorcem:

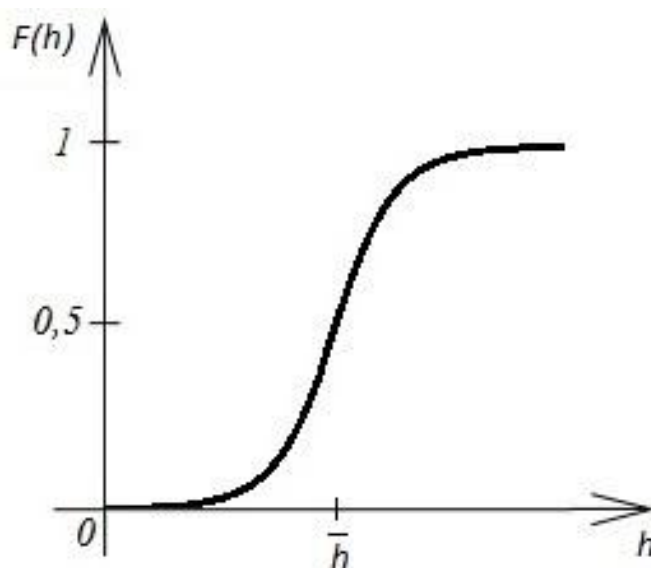
$$p(h) = \frac{1}{\sqrt{2 * \pi * P_h}} * e^{-\frac{1}{2} * \frac{(h - \bar{h})^2}{P_h}} = N(\bar{h}, P_h). \quad (2.12)$$

Hustota pravděpodobnosti této náhodné veličiny bude vypadat následovně:



Obrázek 3 - Hustota pravděpodobnosti normálního rozdělení

Distribuční funkce této náhodné veličiny bude vypadat následovně:



Obrázek 4 – Distribuční funkce normálního rozdělení

Je také důležité říci, že v rovnici (2.12) – hustotě pravděpodobnosti – máme dva parametry. Je to parametr \bar{h} , který je označován jako střední hodnota, a parametr

P_h , kterému se říká variance náhodné veličiny. Střední hodnota a variance je příklad momentů náhodné veličiny. Momenty mohou být využity jako alternativní popis náhodné veličiny. Tyto momenty mohou být buď centrální, nebo necentrální. Momentů centrálních i necentrálních musí být nekonečně mnoho pro kompletní popis náhodné veličiny. Pro příklad si vezmeme náhodnou veličinu H a nejdříve se podíváme, jak je definovaný necentrální moment:

$$E[h^i] = \int_{-\infty}^{+\infty} h^i * p(h) dh, \quad (2.13)$$

kde index i značí, o kolikátý necentrální moment se jedná. Prvnímu necentrálnímu momentu se také říká střední hodnota. Pokud se tedy podíváme na předchozí příklad, kde byla nadefinována náhodná veličina H , kterou lze popsat rovnicí (2.12), parametr \bar{h} bude roven $E[h]$ čili střední hodnotě neboli prvnímu necentrálnímu momentu. Pokud se nyní podíváme na centrální moment a opět použijeme náhodnou veličinu H , jeho definice bude vypadat takto:

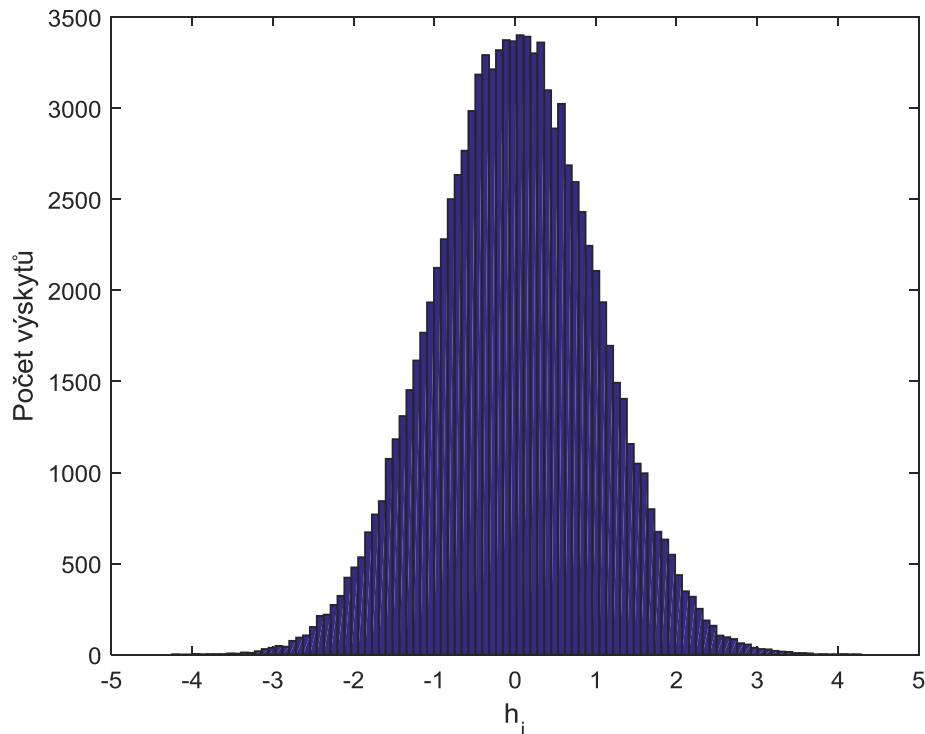
$$E[(h - E[h])^i] = \int_{-\infty}^{+\infty} (h - E[h])^i * p(h) dh, \quad (2.14)$$

kde index i značí, o jaký centrální moment se jedná. Pokud bychom provedli výpočet prvního centrálního momentu, zjistili bychom, že se rovná nule. Zato druhý centrální moment už je nenulový a je označován jako variance nebo rozptyl. Pokud se tedy opět podíváme na předchozí příklad, kde byla nadefinována náhodná veličina H , kterou lze popsat rovnicí (2.12), tak parametr P_h bude roven $E[(h - E[h])^2]$ čili varianci neboli druhému centrálnímu momentu. Pokud popíšeme libovolnou náhodnou veličinu prvním necentrálním momentem a druhým centrálním momentem, používá se zjednodušený výrok, že náhodná veličina je popsána prvními dvěma momenty. Tento popis je ovšem pro libovolnou náhodnou veličinu nedostačující, neboť kdybychom chtěli popsat libovolnou náhodnou veličinu plně pomocí momentů, museli bychom vypočítat nekonečně momentů, což je nerealizovatelné. Na druhou stranu například pro náhodnou veličinu H , která byla popsána vzorcem (2.12) a jež byla gaussovská, první dva momenty plně popisují celou náhodnou veličinu.

Další z možných popisů náhodné veličiny je popis pomocí množiny vzorků či měření. Pro příklad využijeme náhodnou veličinu H a řekněme, že je gaussovská, má střední hodnotu rovnou nule a varianci, která je rovna jedné. Pokud tedy provedeme

sérii měření, abychom získali množinu vzorků, můžeme pomocí této množiny vypočítat různé odhady. Pomocí programu Matlab si tedy vygenerujeme vzorky $\{h_i\}_{i=1}^{100\,000}$, této gaussovské náhodné veličiny H a podíváme se na to, jaké odhady můžeme provést.

Zprvė můžeme data h_i například vykreslit pomocí histogramu, který vypadá následovně:



Obrázek 5 – Histogram realizací normální veličiny

Z tohoto histogramu lze vyčíst, že jeho tvar nejspíše odpovídá hustotě pravděpodobnosti gaussovského rozdělení jako na obrázku 3, a tudíž že i náhodná veličina bude gaussovského rozdělení. Také se dá říci, že nejspíše bude tato náhodná veličina mít střední hodnotu nulovou. Pomocí známých vzorců lze také přímo vypočítat odhad prvních momentů. Vzorec pro nestranný odhad střední hodnoty vypadá následovně:

$$\hat{h} = \widehat{E[h]} = \frac{1}{N} \sum_{i=1}^N h_i, \quad (2.15)$$

kde N značí počet vzorků a index i značí, o které realizace se jedná. V našem případě je N rovno 100 000. Tímto výpočtem získáme tedy odhad, který je roven $-4.4818 \cdot 10^{-4}$, což je velice blízké skutečné střední hodnotě, která je rovna nule.

Že je odhad nestranný, se můžeme jednoduše přesvědčit tím, že dosadíme do definice nestrannosti, která vypadá následovně:

$$E[\widehat{E}[h]] = E\left[\frac{1}{N} * \sum_{i=1}^N h_i\right] = \frac{1}{N} * \sum_{i=1}^N E[h_i] = E[h] = \bar{h}. \quad (2.16)$$

Z tohoto výpočtu lze vidět, že střední hodnota odhadu střední hodnoty je nestranná, neboť se rovná skutečné střední hodnotě.

Je zde důležité také zmínit několik pravidel, jež platí při výpočtu střední hodnoty a vypadají následovně:

$$E[h + d] = E[h] + E[d], \quad (2.17)$$

$$E[c * h] = c * E[h], \quad (2.18)$$

$$E[h + c] = E[h] + c, \quad (2.19)$$

kde d je libovolná náhodná veličina a c je libovolná konstanta.

Nyní si zadefinujme vzorec, kterým se počítá druhý moment, čili varianci na základě střední hodnoty a druhého necentrálního momentu, který vypadá následovně:

$$var(h) = E[h * h^T] - E[h] * E[h]^T = E[(h - E[h]) * (h - E[h])^T], \quad (2.20)$$

kde var značí výpočet variance. Odhad této variance se provede pomocí následujícího vzorce:

$$\widehat{var}(h) = \frac{1}{N} * \sum_{i=1}^N (h_i - E[h]) * (h_i - E[h])^T. \quad (2.21)$$

Nejdříve si ukážeme nestrannost tohoto odhadu jako v předchozím případě, kterou ověříme následujícím výpočtem:

$$\begin{aligned}
E[\widehat{var}(h)] &= E\left[\frac{1}{N} * \sum_{i=1}^N (h_i - E[h]) * (h_i - E[h])^T\right] = \\
&= \frac{1}{N} * \sum_{i=1}^N E[h_i * h_i^T - E[h] * h_i^T - h_i * E[h]^T + E[h] * E[h]^T] = \quad (2.22) \\
&= E[h_i * h_i^T] - 2 * E[h] * E[h]^T + E[h] * E[h]^T = \\
&= E[h * h^T] - E[h] * E[h]^T = var(h) = P_h
\end{aligned}$$

Z tohoto výpočtu je vidět, že se střední hodnota odhadu variance rovná varianci, a tudíž je odhad nestranný. Toto ovšem platí pouze v případě, že budeme znát střední hodnotu skutečného rozdělení.

Pokud totiž do výpočtu odhadu variance dosadíme pouze odhad střední hodnoty, výpočet nestrannosti vyjde následovně:

$$\begin{aligned}
E[\widehat{var}(h)] &= E\left[\frac{1}{N} * \sum_{i=1}^N (h_i - \widehat{E[h]}) * (h_i - \widehat{E[h]})^T\right] = \\
&= \frac{1}{N} * \sum_{i=1}^N E[h_i * h_i^T - \widehat{E[h]} * h_i^T - h_i * \widehat{E[h]}^T + \widehat{E[h]} * \widehat{E[h]}^T] = \\
&= E\left[h * h^T - \left(\frac{1}{N} \sum_{i=1}^N h_i\right) * h^T - h * \left(\frac{1}{N} \sum_{i=1}^N h_i\right)^T + \left(\frac{1}{N} \sum_{i=1}^N h_i\right) * \left(\frac{1}{N} \sum_{i=1}^N h_i\right)^T\right] = \quad (2.23) \\
&= var(h) + E[h] * E[h]^T - \frac{2}{N} * (var(h) + E[h] * E[h]^T + (N - 1) * E[h] * E[h]^T) + \\
&\quad + \frac{1}{N^2} * [(var(h) + E[h] * E[h]^T) * N + E[h] * E[h]^T * (N^2 - N)] = \\
&= var(h) * \frac{(N - 1)}{N} \neq var(h).
\end{aligned}$$

Jak je vidět, pokud neznáme střední hodnotu, nelze jí v rovnici (2.21) nahradit odhadem, neboť poté dává rovnice stranné odhady. Jde to ale napravit tím, že rovnici (2.23) upravíme do následujícího tvaru:

$$\widehat{var}(h) = \frac{1}{N-1} * \sum_{i=1}^N (h_i - \widehat{E}[h]) * (h_i - \widehat{E}[h])^T. \quad (2.24)$$

Takto upravená rovnice dává při nahrazení střední hodnoty jejím odhadem nestranné odhady variance. Pokud tedy pomocí rovnice (2.24) vypočteme kovarianci množiny h , jež je vykreslena na obrázku 5, bude se odhad variance rovnat 1.0005.

Stejně jako u střední hodnoty je důležité zmínit několik vlastností při výpočtu variance, jež jsou vidět na následujících výpočtech:

$$var(h + c) = var(h), \quad (2.25)$$

$$var(c * h) = c^2 * var(h), \quad (2.26)$$

kde c je libovolná konstanta.

Další popis neurčitosti je pomocí kovariance. Je to vlastně hodnota, která nám říká míru lineární závislosti dvou realizací náhodných veličin. Pro příklad využijeme zavedené náhodné proměnné d a h .

Nejdříve si zdefinujeme vzorec, kterým se počítá kovariance, vypadá následovně:

$$cov(d, h) = E[d * h^T] - E[d] * E[h]^T = E[(d - E[d]) * (h - E[h])^T], \quad (2.27)$$

kde cov značí funkci kovariance. Odhad kovariance se vypočte pomocí následujícího vztahu:

$$\widehat{cov}(d, h) = \frac{1}{N} * \sum_{i=1}^N (d_i - E[d]) * (h_i - E[h])^T. \quad (2.28)$$

Rovnice (2.28) dává nestranné odhady stejně jako u rovnice (2.21) pouze v případě, že známe střední hodnoty obou veličin. Pokud bychom měli pouze odhady těchto středních hodnot, vzorec (2.28) se musí upravit, pokud chceme dostávat nestranné odhady, takto:

$$\widehat{cov}(d, h) = \frac{1}{N-1} * \sum_{i=1}^N (d_i - \widehat{E}[d]) * (h_i - \widehat{E}[h])^T. \quad (2.29)$$

Stejně jako u střední hodnoty a variance existují určitá pravidla při výpočtu kovariance a vypadají následovně:

$$\begin{aligned} cov(h, b * h + c) &= E[h * (b * h + c)^T] - E[h] * E[b * h + c]^T = \\ &= b * E[h * (h)^T] + c - b * E[h] * E[(h)]^T - c = b * var(h), \end{aligned} \quad (2.30)$$

kde c a b jsou libovolné konstanty.

Dosud jsme uvažovali jednotlivé realizace h_i , které byly navzájem nezávislé a kde index i značil pouze, o kterou realizaci jde. Lze ale například uvažovat následující proces:

$$h_{i+1} = \lambda * h_i + \xi_i, \quad (2.31)$$

kde ξ je náhodná veličina popsaná jako $N(0, R_\xi)$ a konstantní parametr λ je menší v absolutní hodnotě než jedna. U takového procesu dolní index značí, v jakém časovém okamžiku byla realizace změřena. Lze tedy využít například rovnici (2.29), u které bychom nepočítali s dvěma realizacemi, ale použili bychom jednu realizaci h_i jak je vidět v následujícím výpočtu:

$$\widehat{cov}(h_j) = \frac{1}{N-j-1} * \sum_{i=1}^N (h_{i-j} - \widehat{E}[h]) * (h_i - \widehat{E}[h])^T, \quad (2.32)$$

kde parametr j umožňuje časový posun realizací z procesu popsaného rovnicí (2.31). Tímto výpočtem lze vypočítat kovarianci jedné náhodné veličiny posunuté o libovolný časový okamžik. Obvykle je rozumné uvažovat velké N , při kterém se proces (2.31) ustálí tak, aby počáteční podmínky neovlivňovaly správnost výpočtu rovnice (2.32).

2.4. Stavový stochastický popis systému

V kapitole 2.2 jsme si ukázali, jak popsat systém deterministicky. Tento popis je však často nevyhovující, zejména proto, že takto popsany systém neodpovídá zkoumané realitě. Těmito důvody jsou například nepřesnost v deterministické části, která vznikne při matematickém modelování, neboť realitu nelze nikdy popsat dokonale. Dalším důvodem jsou také poruchy, které působí na stav systému a nedají se nijak předpovídat, nebo například určitá nepřesnost měřících přístrojů či senzorů měřící výstup systému. To je několik hlavních důvodů k zavedení neurčitosti do deterministického modelu a tím vytvoření stochastického modelu.

Stochastický model by tedy vypadal následovně, pokud bychom pro zjednodušení zvolili vektor stavu jako x , vektor vstupů u , vektor měření y , vektory poruch w a v a matice budou značeny A , B a C . Takto zvolený proces by vypadal následovně:

$$x_{k+1} = A * x_k + B * u_k + w_k \quad (2.33)$$

$$y_k = C * x_k + v_k. \quad (2.34)$$

V rovnici (2.33) přibyl vektor w . Tento vektor je bílá porucha $w \in \mathbb{R}^{n \times 1}$, kterou lze reprezentovat třemi způsoby. Zaprvé jako nepřesnost systému, které jsme se mohli dopustit při modelování reálného procesu, zadruhé jako náhodnou poruchu působící na systém, kterou nelze měřit a tím jí započítat jako korekci do systému, a zatřetí jako neurčitost, která vzniká při generování řízení. Vektor w bude mít gaussovské rozdělení s nulovou střední hodnotou a kovarianční matice bude symetrická pozitivně definitní matice $Q \in \mathbb{R}^{n \times n}$. Poruchu lze popsat jako $w \sim N(0, Q)$.

V rovnici (2.34) přibyl vektor v . Tento vektor je bílá porucha $v \in \mathbb{R}^{p \times 1}$, kterou lze prezentovat jako poruchu, která vzniká jednak nepřesností systému, jednak nepřesností senzorů měřících výstup systému. Vektor v bude mít gaussovské rozdělení s nulovou střední hodnotou a kovarianční matice bude symetrická pozitivní definitní matice $R \in \mathbb{R}^{p \times p}$. Poruchu lze popsat jako $v \sim N(0, R)$.

3. Odhad stavu stochastického systému

Tato kapitola se zabývá odhadem stavu diskretního dynamického systému pomocí Kalmanova filtru a jeho předpoklady a vlastnostmi.

Pro následující text byly zdrojem [20] a [30].

3.1. Kalmanův filtr

Pro odhad stavu stochastického systému se dnes převážně využívá Kalmanova filtru, dále značený už jen KF. Tento rekurzivní lineární filtr generuje nestranné odhady stavu ve smyslu podmíněné střední hodnoty neznámého dynamického systému ze zašuměných dat změřených v diskretních časových okamžicích. Tento filtr je velice používán například v oblastech automatických řídicích, komunikačních, monitorovacích či navigačních systémech.

Tento filtr byl navržen v 60. letech a obsahuje dvě části – filtraci a predikci. KF vychází z lineárního diskretního stavového popisu systému, jako jsou rovnice (2.33) a (2.34). Je také důležité říci, že tyto rovnice musí obsahovat poruchy w a v , které jsou bílé, jak už bylo nadefinováno, ale navíc musí být tyto poruchy navzájem nekorelované, jejich vzájemná kovariance tak musí být nulová.

3.2. Kalmanův filtr – filtrace

Rovnice KF vychází ze vztahu pro výpočet optimálního aposteriorního odhadu pomocí lineárního filtru, který vypadá následovně:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + P_{xy_{k|k-1}} * P_{y_{k|k-1}}^{-1} * (y_k - \hat{y}_{k|k-1}), \quad (3.1)$$

kde \hat{x} je odhad stavu systému, $P_{xy_{k|k-1}}$, dále už jen P_{xy} , je prediktivní kovariance chyby odhadu stavu a odhadu měření, $P_{y_{k|k-1}}$, dále už jen P_y , je kovariance chyby odhadu měření, y je měření a \hat{y} je odhad měření. Index například u $\hat{x}_{k|k}$ značí, že tento odhad stavu je v čase k , jak značí první písmeno v indexu, a druhé písmeno značí, že bylo použito měření do času k . Proto je rovnice (3.1) filtrační, neboť jsme měli měření i v čase k . Součin kovariance P_{xy} a P_y^{-1} se také značí jako Kalmanův zisk $L \in \mathbb{R}^{n \times p}$, dále značený už jen KZ.

Při pohledu na rovnici (3.1) lze říci, že pro odhad stavu v čase k musíme znát predikci z času $k - 1$ do času k a dále pak měření v čase k a predikci měření z času

$k - 1$ do času k . Rozdíl měření v čase k a jeho predikcí nám vlastně říká rozdíl chyby, jaké jsme se dopustili při výpočtu predikce. Tato chyba je zpočátku zapříčiněná počáteční volbou stavu v čase 0, kdy jsme ještě neměli měření a kdy stav nešlo odhadnout. Dalším důvodem této chyby je porucha, která působí jak na stav, tak i na měření systému, především kvůli tomuto je zaveden a používán KF, aby se tato porucha co nejvíce eliminovala. Tato chyba se násobí KZ L tak, aby korekce, kterou připočteme k predikci stavu, byla optimální, aby se odhad stavu v čase k blížil co nejvíce ke skutečné hodnotě. Je také vidět, že pokud by neexistovala žádná lineární závislost stavu x a měření y , byla by kovariance P_{xy} rovna nule a tím by měření nijak neovlivnilo predikci stavu do času k , a tak by se predikce stavu do času k rovnala filtraci v čase k . Pokud by naopak kovariance P_{xy} byla velká, měření by ovlivnilo predikci hodně. Další kovariancí, která ovlivňuje tuto chybu, je P_y . Tato kovariance nám vlastně říká, jak moc měření y kolísá kvadratickou odchylkou od její střední hodnoty. Toto kolísání způsobuje porucha ovlivňující měření, pokud by tedy výstup hodně kolísal, byl by tedy zatížen velkou poruchou, tak by se měření opět neuplatnilo ve filtraci v čase k . Stejně tak, pokud by měření bylo konstantní, kovariance P_y by bylo hodně malá a měření by se uplatnilo hodně ve filtraci v čase k .

Pro použití rovnice (3.1), ale musíme nejdříve znát odhad měření, který vypočteme dle následujícího vztahu:

$$\hat{y}_{k|k-1} = E[y_k] = E[C * x_{k|k-1} + v_k] = C * \hat{x}_{k|k-1}. \quad (3.2)$$

Dále pak musíme vypočítat kovariance z rovnice (3.1) která vycházejí následovně:

$$\begin{aligned} P_y &= E[y_k * y_k^T] - \hat{y}_{k|k-1} * \hat{y}_{k|k-1}^T = \\ &= E[(C * x_k + v_k) * (C * x_k + v_k)^T] - C * \hat{x}_{k|k-1} * \hat{x}_{k|k-1}^T * C^T = \\ &= C * (E[x_k * x_k^T] - \hat{x}_{k|k-1} * \hat{x}_{k|k-1}^T) * C^T + E[v_k * v_k^T] = \\ &= C * P_{k|k-1} * C^T + R, \end{aligned} \quad (3.3)$$

$$\begin{aligned}
P_{xy} &= E[x_k * y_k^T] - \hat{x}_{k|k-1} * \hat{y}_{k|k-1}^T = \\
&= E[x_k * (C * x_k + v_k)^T] - \hat{x}_{k|k-1} * \hat{x}_{k|k-1} * C^T = \\
&= (E[x_k * x_k^T] - \hat{x}_{k|k-1} * \hat{x}_{k|k-1}) * C^T = P_{k|k-1} * C^T.
\end{aligned} \tag{3.4}$$

Pokud tyto vypočtené vztahy (3.2), (3.3) a (3.4) dosadíme do prvotní rovnice (3.1), vztah bude vypadat následovně:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + P_{k|k-1} * C^T * (C * P_{k|k-1} * C^T + R)^{-1} * (y_k - C * \hat{x}_{k|k-1}). \tag{3.5}$$

Tento vztah se také často přepisuje do následujícího vztahu:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + L_k * (y_k - C * \hat{x}_{k|k-1}), \tag{3.6}$$

kde KZ L je definován takto:

$$L_k = P_{k|k-1} * C^T * (C * P_{k|k-1} * C^T + R)^{-1}. \tag{3.7}$$

Tento zisk se v průběhu odhadu stavu mění, až nakonec dokonverguje k optimální hodnotě, která způsobí, že odhad stavu bude mít minimální možnou odchylku od skutečného stavu, neboť poruchy na stavu a měření jsou nepředvídatelné a tudíž je nelze zcela eliminovat.

Také je důležité upozornit, že závorka $(y_k - C * \hat{x}_{k|k-1})$ je nazývána inovace, která v jednotlivých časových okamžicích tvoří společně takzvanou inovační posloupnost. Tato posloupnost by měla z převážné části, pokud tedy dojde k optimálnímu nastavení KZ L , být utvářena poruchami působícími na systém a měření. Z toho také vyplývá, že tato inovační posloupnost by měla být, jako jsou poruchy, bílá. Bělost inovační posloupnosti se často využívá jako kontrola kvality KF. Pokud by byla například porucha na stavu nulová, byla by inovační posloupnost tvořená pouze poruchou měření. Naopak pokud by porucha na měření byla nulová, byla by inovační posloupnost tvořená poruchou stavu, která by se na výstupu a tím i do inovační posloupnosti dostávala skrze stav systému.

Odhaduje se také kovariance chyby odhadu stavu, kterou značíme $P_{k|k}$. Její hodnota nám říká důvěryhodnost našeho odhadu filtrovaného stavu, který je vypočten v rovnici (3.1).

Kovariance se vypočte klasickým způsobem, jak je vidět v následujícím výpočtu:

$$\begin{aligned}
P_{k|k} &= E[x_k * x_k^T] - \hat{x}_{k|k} * \hat{x}_{k|k}^T = \\
&= E[(x_k - \hat{x}_{k|k}) * (x_k - \hat{x}_{k|k})^T] = \\
&= E\left[(x_k - \hat{x}_{k|k-1} - P_{xy} * P_y^{-1} * (y_k - \hat{y}_{k|k-1})) * \right. \\
&\quad \left. * (x_k - \hat{x}_{k|k-1} - P_{xy} * P_y^{-1} * (y_k - \hat{y}_{k|k-1}))^T\right] = \\
&= E[(x_k - \hat{x}_{k|k-1}) * (x_k - \hat{x}_{k|k-1})^T] - \\
&\quad - E[(x_k - \hat{x}_{k|k-1}) * (y_k - \hat{y}_{k|k-1})^T] * P_y^{-1T} * P_{xy}^T - \\
&\quad - P_{xy} * P_y^{-1} * E[(y_k - \hat{y}_{k|k-1}) * (x_k - \hat{x}_{k|k-1})^T] + \tag{3.8} \\
&\quad + P_{xy} * P_y^{-1} * E[(y_k - \hat{y}_{k|k-1}) * (y_k - \hat{y}_{k|k-1})^T] * P_y^{-1T} * P_{xy}^T = \\
&= P_{k|k-1} - P_{xy} * P_y^{-1T} * P_{xy}^T - P_{xy} * P_y^{-1} * P_{xy}^T + \\
&\quad + P_{xy} * P_y^{-1} * P_y * P_y^{-T} * P_{xy}^T = \\
&= P_{k|k-1} - P_{xy} * P_y^{-1} * P_{xy}^T = \\
&= P_{k|k-1} - P_{k|k-1} * C^T * (C * P_{k|k-1} * C^T + R)^{-1} * C * P_{k|k-1}^T = \\
&= (I_{n \times n} - L_k * C) * P_{k|k-1},
\end{aligned}$$

kde matice $I_{n \times n}$ je jednotková matice dimenze $n \times n$.

Pokud bychom si zavedli odchylku predikce stavu v čase k a skutečného stavu v čase k , jak je vidět na následujícím vztahu:

$$\tilde{x}_k = x_k - \hat{x}_{k|k-1}, \tag{3.9}$$

byla by střední hodnota této chyby, dána následujícím vztahem:

$$E[\tilde{x}_k] = E[(x_k - \hat{x}_{k|k-1})] = 0. \tag{3.10}$$

Je dobré si také všimnout z prvního řádku rovnice (3.8), že kovariance stavu je vlastně totožná s kovariancí chyby zavedené v rovnici (3.9), jak je vidět v následujícím vztahu:

$$P_{k|k} = E[x_k * x_k^T] - \hat{x}_{k|k-1} * \hat{x}_{k|k-1}^T = E[\tilde{x}_k * \tilde{x}_k^T]. \quad (3.11)$$

Pokud by byla tedy kovariance chyby \tilde{x}_k malá, byla by tím pádem také malá i kovariance stavu $P_{k|k}$, což by znamenalo, že by odhad stavu v rovnici (3.1) měl velkou důvěryhodnost.

3.3. Kalmanův filtr – predikce

Kalmanova rovnice pro výpočet predikce odhadu stavu vypadá následovně:

$$\hat{x}_{k+1|k} = E[A * x_k + B * u_k + w_k] = A * \hat{x}_{k|k} + B * u_k. \quad (3.12)$$

Je zřejmé, že jde o predikci z času k do času $k + 1$.

Kalmanova rovnice pro výpočet predikce kovariance odhadu stavu $P_{k+1|k}$ je vidět v následujícím výpočtu:

$$\begin{aligned} P_{k+1|k} &= E[x_{k+1} * x_{k+1}^T] - \hat{x}_{k+1|k} * \hat{x}_{k+1|k}^T = \\ &= A * (E[x_k * x_k^T] - \hat{x}_{k|k} * \hat{x}_{k|k}^T) * A^T + E[w_k * w_k^T] = \\ &= A * P_{k|k} * A^T + Q. \end{aligned} \quad (3.13)$$

Její hodnota nám říká důvěryhodnost naší predikce, kterou jsme provedli v rovnici (3.12).

Pokud bychom do této rovnice dosadili rovnici filtrace (3.8), do které byla dosazena rovnice KZ (3.7), získáme následující vztah:

$$\begin{aligned} P_{k+1|k} &= A * P_{k|k-1} * A^T - \\ &- A * P_{k|k-1} * C^T * (C * P_{k|k-1} * C^T + R)^{-1} * C * P_{k|k-1} * A^T + Q. \end{aligned} \quad (3.14)$$

Této rovnici se říká Riccatiho rovnice, která má jediné ustálené pozitivně semidefinitní řešení za předpokladu stability matice A , symetričnosti pozitivně definitních matic Q a R , detekovatelnosti dvojice matic A a C a stabilizovatelnosti dvojice matic A a Γ , kde $Q = \Gamma * \Gamma^T$, viz [30]. To znamená, že se bude predikce

kovariance stavu v čase k do $k + I$ rovnat predikci v čase $k - I$ do času k . Poté bude mít rovnice řešení P , které lze vypočítat například pomocí příkazu `dare(AT, CT, Q, R)` programu Matlab. Pomocí tohoto ustáleného řešení P lze například vypočítat optimální KZ v rovnici (3.7) následovně:

$$L_{\text{optimální}} = P * C^T * (C * P * C^T + R)^{-1}. \quad (3.15)$$

Je také důležité si uvědomit, že prediktivní kovariance stavu z rovnice (3.13), ale i filtrační kovariance stavu z rovnice (3.8) lze vypočítat dopředu, bez znalosti jakéhokoliv měření.

4. Motivace a cíl práce

Tato kapitola se zabývá motivací a specifikací cílů této diplomové práce.

4.1 Předpoklady pro Kalmanův filtr

Při návrhu a implementaci KF existuje vždy několik kroků a předpokladů, které musí být splněny.

Jedním z těchto předpokladů je například znalost modelu tj. matic A , B a C . To ovšem nebývá ve většině případů limitující, neboť, jak bylo diskutováno v kapitole 2.1, tyto matice lze získat za pomoci fyzikálních zákonů či experimentálních metod identifikace za předpokladu nezávislosti na chybách při měření experimentu.

Dalším předpokladem je znalost kovariančních matic poruch Q a R . Tato znalost ovšem není nijak jednoduchá. Například při ideálních podmínkách kdybychom věděli, že hodnota stavu je kupříkladu neměnná vzdálenost od zdi, která by byla měřena ultrazvukovým měřičem vzdálenosti. Díky tomuto měření by matice dynamiky byla jednotková, stav by se v čase neměnil. Na tento stav bychom tedy neuvažovali ani poruchu prezentovanou kovarianční maticí Q . Poté by stačilo pouze naměřit dostatečně velkou množinu dat, která by reprezentovala měření y . Pokud bychom vypočítali kovarianci tohoto měření y , získali bychom tím dokonce nestranný odhad kovarianční matice R , která by byla totožná s kovariancí měření. Toto je ovšem velice specifický případ, u kterého lze zaručit neměnnost stavu. Například při měření vzdálenosti vrtulníku od země by takovéto zaručení neměnného stavu jako vzdálenosti nešlo vůbec realizovat. Těmito příklady byla ukázána určitá možnost, ale spíše potíže, které by vedly k identifikaci matice R . Pro zjištění kovarianční matice Q by nám stačilo znát přesné měření a tím i přesný stav. Díky tomu bychom mohli identifikovat kovarianční matici Q . Tento předpoklad je ale neproveditelný, neboť neexistuje senzor, který měří bez chyb.

Zjištění obou matic, tj. kdyby na systém působily obě poruchy najednou, je velice obtížné. Existuje několik metod, které mají ale mnohdy velice omezující předpoklady. Také dávají odhady, u nichž nelze prokázat nestrannost, neboť jsou založené na různých aproximacích.

4.2 Odhad kovariančních matic poruch na základě chyby odhadu lineárního filtru

Metoda ACLS navržená v [3] a dále rozšířená v [4],[5] a [21] je významná z mnoha důvodů. Jedním důvodem je, že pro systém s bílými poruchami působícími na systém a měření poskytuje tato metoda nestranné odhady kovariančních matic těchto poruch. Dalším důvodem je například to, že metoda je analyticky odvozená.

Metoda využívá lineární filtr, u kterého je volba zisku ponechána na uživateli, a tak není ani optimální. Díky této úpravě, znalosti systému a vypočítání odhadů kovariancí inovační posloupnosti lze pomocí identifikační metody nejmenších čtverců odhadnout kovarianční matice poruch působící na systém a měření, díky čemuž lze plně využít například KF pro odhad neznámého stavu či parametrů systému.

Jako každá identifikační metoda má i tato své předpoklady. Jedním z těchto předpokladů jsou například vlastnosti poruch působící na systém. Tyto poruchy musí být bílé, navzájem nekorelované a také časově nekorelované. Dalším předpokladem je například to, že systém musí být t -invariantní.

Také tato metoda má i svá omezení či nežádoucí vlastnosti. Mezi tyto vlastnosti patří za určitých podmínek omezený počet odhadnutelných prvků kovariančních matic, z kolika rovnic kovariance inovační posloupnosti se budou neznámé parametry pomocí metody nejmenších čtverců odhadovat nebo například otázka, jak zvolit konstantní zisk L tak, aby byla kvalita odhadnutých kovariančních matic byla co největší a aby byla metoda vůbec stabilní.

4.3 Cíl práce

Prvním cílem této práce je analyzovat a následně najít vhodnou volbu konstantního zisku L v lineárním prediktoru. Jednoduchými simulacemi se dá ukázat, že volba tohoto zisku ovlivňuje kvalitu odhadu kovariančních matic poruch Q a R . Proto také bude v této práci jedním z cílů analyzovat vhodnou volbu tohoto zisku.

Druhým cílem bude implementovat tuto metodu v softwarovém programu Matlab, kde budou metoda a její vlastnosti testovány.

Posledním cílem bude ukázat, že tato metoda se dá vhodnou volbou konstantního zisku L rozšířit i pro případy, kdy na měření bude působit mnohem komplexnější porucha. Struktura této komplexní poruchy byla navržena tak, aby odpovídala poruchám, které působí na inercionální senzory, které se využívají dnes v mnoha aplikacích sahajících od mobilních telefonů až po letectví.

5. Autokovarianční metoda pro odhad kovariančních matic poruch LS

Tato kapitola se zabývá metodou ACLS jejím odvozením, vlastnostmi a také specifikací omezení či nedostatků.

5.1 Odvození metody

Tato metoda je navržena pro systém, který je popsán rovnicemi (2.28), (2.29) a také diskutována v [3] – [5] a [21]. Pro tento systém lze použít rovnici Kalmanova filtru (3.6) a (3.12). Pokud bychom dosadili rovnici filtrace (3.6) do rovnice predikce (3.12), získáme tím rovnici, která vypadá následovně:

$$\hat{x}_{k+1|k} = A * \hat{x}_{k|k-1} + A * L_k * (y_k - C * \hat{x}_{k|k-1}) + B * u_k. \quad (5.1)$$

Pokud bychom navíc za měření v čase k dosadili rovnici měření, získáme tento vztah:

$$\hat{x}_{k+1|k} = A * \hat{x}_{k|k-1} + A * L_k * (C * x_k + v_k - C * \hat{x}_{k|k-1}) + B * u_k. \quad (5.2)$$

Dále bychom si nadefinovali chybu prediktivního odhadu stavu, který vypadá následovně:

$$\varepsilon_k = x_k - \hat{x}_{k|k-1}, \quad (5.3)$$

či o jeden krok posunutou:

$$\varepsilon_{k+1} = x_{k+1} - \hat{x}_{k+1|k}. \quad (5.4)$$

Tato chyba odpovídá odchylce predikovaného stavu v čase $k + 1$ a skutečného stavu v čase $k + 1$. Pokud bychom do tohoto vztahu (5.4) dosadili rovnici (5.2), získáme tento vztah:

$$\varepsilon_{k+1} = x_{k+1} - A * \hat{x}_{k|k-1} - A * L_k * (C * x_k + v_k - C * \hat{x}_{k|k-1}) - B * u_k. \quad (5.5)$$

Dále bychom za stav v čase $k + 1$ dosadili stav v čase k z rovnice stavu (2.33) a tím bychom získali tento vztah:

$$\varepsilon_{k+1} = A * x_k + B * u_k + w_k - A * \hat{x}_{k|k-1} - A * L_k * (C * x_k + v_k - C * \hat{x}_{k|k-1}) - B * u_k. \quad (5.6)$$

Další úpravou lze rovnici (5.6) převést na tento vztah:

$$\varepsilon_{k+1} = A * (x_k - \hat{x}_{k|k-1}) - A * L_k * C(x_k - \hat{x}_{k|k-1}) + w_k - A * L_k * v_k. \quad (5.7)$$

Ze znalosti zavedeného stavu (5.3) lze rovnici (5.7) převést na tento vztah:

$$\varepsilon_{k+1} = (A - A * L_k * C) * \varepsilon_k + [I_{n \times n} \quad -A * L_k] * \begin{bmatrix} w_k \\ v_k \end{bmatrix}. \quad (5.8)$$

Tato metoda také zavádí konstantní zisk L , který nahradí Kalmanův zisk L_k , který byl v každém kroku přepočítáván, dokud se neustálil na optimální hodnotě. Tento výpočet optimálního Kalmanova zisku ovšem potřebuje znalost kovarianční matice R , která je přímo součástí při výpočtu Kalmanova zisku, a také znalost kovarianční matice Q , která se do výpočtu dostává skrze výpočet predikce kovariance stavu x , jenž je odhadován pomocí Kalmanova filtru. Díky volbě konstantního zisku lze ale tuto neznalost matic R a Q určitým způsobem obejít. To však nese i své záporné stránky, mezi ně patří skutečnost, že takto zvolený zisk není optimální, a tak i z Kalmanova filtru nelze získávat nejlepší odhady stavu x . To ale této metodě nevadí, neboť jí o správný odhad stavu ani nejde.

Dále nadefinujeme nové matice \bar{A} , \bar{G} a vektor \bar{w}_k , které odpovídají následujícím vztahům:

$$\bar{A} = (A - A * L * C), \quad (5.9)$$

$$\bar{G} = [I_{n \times n} \quad -A * L], \quad (5.10)$$

$$\bar{w}_k = \begin{bmatrix} w_k \\ v_k \end{bmatrix}. \quad (5.11)$$

Díky tomu lze rovnici (5.8) přepsat do následující kompaktní formy:

$$\varepsilon_{k+1} = \bar{A} * \varepsilon_k + \bar{G} * \bar{w}_k, \quad (5.12)$$

$$\bar{A} \in \mathbb{R}^{n \times n}; \bar{G} \in \mathbb{R}^{n \times (n+p)}; \bar{w} \in \mathbb{R}^{(n+p) \times 1}, \quad (5.13)$$

kde rovnice (5.12) popisuje rovnici dynamiky chyby odhadu stavu rovnice (5.3). Rovnice (5.13) ukazuje příslušné dimenze matic a vektoru \bar{w} , který je tvořen chybou w a v a jehož popis přes první dva momenty vypadá takto:

$$\bar{w} \sim N\left(0, \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\right) = N(0, \bar{Q}). \quad (5.14)$$

Abychom získali popis celého chybového modelu, musíme také zavést výstupní rovnici, která vypadá následovně:

$$z_k = y_k - \hat{y}_{k|k-1} = C * x_k + v_k - C * \hat{x}_{k|k-1} = C * \varepsilon_k + v_k. \quad (5.15)$$

Důležité je, že tento výstup musí být měřitelný, a proto byl také zvolen jako inovační posloupnost, kterou lze získat z rovnic Kalmanova filtru (3.6) a (3.12), přičemž v rovnici (3.6) je, jak už bylo psáno, Kalmanův zisk nahrazen konstantním ziskem.

Z rovnic (5.12) a (5.15) máme tedy zavedený chybový model, jehož výstup je inovační posloupnost. Pro stabilitu tohoto systému z teorie stability pro diskrétní dynamické systémy je důležité, aby systém měl všechny póly systému, v našem případě všechny vlastní čísla matice \bar{A} , stabilní, tj. aby byly menší než jedna v absolutní vzdálenosti od počátku. To také znamená, že volba konstantního zisku L v rovnici (5.9) má svá omezení.

Následujícím výpočtem si ukážeme, že veličina ε má střední hodnotu nulovou:

$$E[\varepsilon_{k+1}] = E[\bar{A} * \varepsilon_k + \bar{G} * \bar{w}_k] = \bar{A} * E[\varepsilon_k] + \bar{G} * E[\bar{w}_k] = \bar{A} * E[\varepsilon_k] = 0, \quad (5.16)$$

neboť střední hodnota poruch $E[\bar{w}_k]$ je rovna nule a za předpokladu stability matice \bar{A} se střední hodnota chyby prediktivního odhadu stavu ustálí, tudíž bude platit $E[\varepsilon_{k+1}] = E[\varepsilon_k]$, tím pádem se střední hodnota chyby prediktivního odhadu bude rovnat nule.

Jako další se nám bude hodit kovariance stavu ε , kterou vypočítáme následujícím výpočtem:

$$\begin{aligned}
cov[\varepsilon_{k+1}] &= E[(\bar{A} * \varepsilon_k + \bar{G} * \bar{w}_k) * (\bar{A} * \varepsilon_k + \bar{G} * \bar{w}_k)^T] = \\
&= E[\bar{A} * \varepsilon_k * \varepsilon_k^T * \bar{A}^T + \bar{G} * \bar{w}_k * \bar{w}_k^T * \bar{G}^T] = \\
&= \bar{A} * cov[\varepsilon_k] * \bar{A}^T + \bar{G} * \bar{Q} * \bar{G}^T.
\end{aligned} \tag{5.17}$$

Pokud bude, jak už bylo řečeno, matice \bar{A} stabilní, bude mít tato Ljapunovova rovnice ustálené řešení P^- . Toto ustálené řešení se vypočítá následujícím výpočtem:

$$cov[\varepsilon] - \bar{A} * cov[\varepsilon] * \bar{A}^T = \bar{G} * \bar{Q} * \bar{G}^T. \tag{5.18}$$

Na tuto rovnici použijeme vektorový operátor vec (viz Příloha), který bude značen pro jednoduchost dolním indexem S . Rovnici lze dále upravit následovně:

$$\begin{aligned}
(I_{n^2 \times n^2} - \bar{A} \otimes \bar{A}) * cov[\varepsilon]_S &= (\bar{G} * \bar{Q} * \bar{G}^T)_S \\
cov[\varepsilon]_S &= (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * (\bar{G} \otimes \bar{G}) * \bar{Q}_S = P^-_S,
\end{aligned} \tag{5.19}$$

kde \otimes značí Kroneckerův součin (viz Příloha).

Dále si ukážeme hlavní podstatu této metody, což je výpočet kovariance inovační posloupnosti neboli také výstupu modelu (5.15). Tento výpočet je založen na kovarianci výstupu v jednom časovém okamžiku. Jelikož víme, že inovační posloupnost má nulovou střední hodnotu, výpočet kovariance vypadá tedy následovně:

$$\begin{aligned}
E[z_k * z_k^T] &= E[(C * \varepsilon_k + v_k) * (C * \varepsilon_k + v_k)^T] = \\
&= E[C * \varepsilon_k * \varepsilon_k^T * C^T + v_k * v_k^T] = \\
&= C * P^- * C^T + R.
\end{aligned} \tag{5.20}$$

Další kovariance bude vypočtena pro výstup (5.15) posunutý o jeden časový index:

$$\begin{aligned}
E[z_{k+1} * z_k^T] &= E[(C * \varepsilon_{k+1} + v_{k+1}) * (C * \varepsilon_k + v_k)^T] = \\
&= E[(C * \bar{A} * \varepsilon_k + C * \bar{G} * \bar{w}_k + v_{k+1}) * (C * \varepsilon_k + v_k)^T] = \\
&= E[C * \bar{A} * \varepsilon_k * \varepsilon_k^T * C^T + C * \bar{G} * \bar{w}_k * v_k^T] = \\
&= C * \bar{A} * P^- * C^T + C * \begin{bmatrix} I_{n \times n} & -A * L \end{bmatrix} * \begin{bmatrix} 0 \\ R \end{bmatrix} = \\
&= C * \bar{A} * P^- * C^T - C * A * L * R.
\end{aligned} \tag{5.21}$$

A takto bychom mohli pokračovat při výpočtu kovariance pro různé posunutí výstupu (5.15):

$$E[z_{k+j} * z_k^T] = C * \bar{A}^j * P^- * C^T - C * \bar{A}^{j-1} * A * L * R, \tag{5.22}$$

kde index $j = 0, \dots, O$ značí hodnotu posuvu jednoho výstupu od druhého, pro které je výpočet kovariance prováděn.

V těchto rovnicích ovšem figurují neznámé kovarianční matice R a Q . Matice Q se ukrývá ve výpočtu ustálené hodnoty kovariance v rovnici (5.19). Tohoto problému se zbavíme, pokud do rovnice (5.20) dosadíme ustálenou hodnotu kovariance z rovnice (5.19) a upravíme ji, jak je vidět v následujícím výpočtu:

$$\begin{aligned}
E[z_k * z_k^T]_S &= (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * (\bar{G} \otimes \bar{G}) * \bar{Q}_S + R_S = \\
&= (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * (\bar{G} * \bar{Q} * \bar{G}^T)_S + R_S = \\
&= (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * \left(\begin{bmatrix} I_{n \times n} & -A * L \end{bmatrix} * \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} * \begin{bmatrix} I_{n \times n} \\ -A * L \end{bmatrix} \right)_S + R_S = \\
&= (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * (Q + A * L * R * L^T * A^T)_S + R_S = \\
&= (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * Q_S + \\
&+ (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * ((A * L) \otimes (A * L)) * R_S + R_S = \\
&= (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * Q_S + \\
&+ ((C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * ((A * L) \otimes (A * L)) + I_{p^2 \times p^2}) * R_S.
\end{aligned} \tag{5.23}$$

V této jedné rovnici máme na pravé straně dvě neznámé – vektory Q_S a R_S . Abychom mohli vypočítat tyto neznámé, musí být minimálně počet rovnic roven počtu neznámých v rovnicích. Proto využijeme rovnici (5.21), do které také dosadíme ustálenou kovarianci z rovnice (5.19) a upravíme ji, jak uvidíme v následujícím výpočtu:

$$\begin{aligned}
E[z_k * z^T_{k-1}]_S &= (C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * (\bar{G} \otimes \bar{G}) * \bar{Q}_S + \\
&\quad + (I_{p \times p} \otimes (C * A * L)) * R_S = \\
&= (C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * (\bar{G} * \bar{Q} * \bar{G}^T)_S + (I_{p \times p} \otimes (C * A * L)) * R_S = \\
&= (C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * \left([I_n \times n \quad -A * L] * \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} * \begin{bmatrix} I_n \times n \\ -A * L \end{bmatrix} \right)_S + \\
&\quad + (I_{p \times p} \otimes (C * A * L)) * R_S = \\
&= (C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * (Q + A * L * R * L^T * A^T)_S + \tag{5.24} \\
&\quad + (I_{p \times p} \otimes (C * A * L)) * R_S = \\
&= (C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * Q_S + \\
&\quad + (C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * ((A * L) \otimes (A * L)) * R_S + \\
&\quad + (I_{p \times p} \otimes (C * A * L)) * R_S = \\
&= (C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * Q_S + \\
&+ \left((C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * ((A * L) \otimes (A * L)) + (I_{p \times p} \otimes (C * A * L)) \right) * R_S.
\end{aligned}$$

Pro další rovnice s neznámou Q_S a R_S lze stejně jako v předchozích případech dosadit ustálený stav P^- . Rovnice (5.22) pro libovolný časový posun vypadá následovně:

$$\begin{aligned}
E[z_k * z^T_{k-j}]_S &= (C \otimes (C * \bar{A}^j)) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * Q_S + \\
&\tag{5.25} \\
&+ \left((C \otimes C * \bar{A}^j) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * ((A * L) \otimes (A * L)) + (I_{p \times p} \otimes (C * \bar{A}^{j-1} * A * L)) \right) * R_S.
\end{aligned}$$

Všechny rovnice jako (5.23), (5.24) a (5.25) lze nyní přepsat do maticového tvaru, jenž vede k následující rovnici:

$$\mathcal{A} * \theta(Q, R) = \mathcal{A} * \begin{bmatrix} Q_S \\ R_S \end{bmatrix} = b, \quad (5.26)$$

kde matice \mathcal{A} , kterou lze rozepsat do dvou matic, která je dána následující rovnicí:

$$\mathcal{A} = [\mathcal{A}_1 \quad \mathcal{A}_2], \quad (5.27)$$

kde \mathcal{A}_1 reprezentuje známou matici, která se násobí s neznámou Q_S , a matice \mathcal{A}_2 prezentuje matici, jež se násobí s neznámou R_S . Podoba těchto matic je vidět v následujících rovnicích:

$$\mathcal{A}_1 = \begin{bmatrix} (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} \\ (C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} \\ \vdots \\ (C \otimes C * \bar{A}^{O-1}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} \end{bmatrix}, \quad (5.28)$$

$$\mathcal{A}_2 = \begin{bmatrix} (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * ((A * L) \otimes (A * L)) + I_{p^2 \times p^2} \\ (C \otimes C * \bar{A}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * ((A * L) \otimes (A * L)) + (I_{p \times p} \otimes (C * A * L)) \\ \vdots \\ (C \otimes C * \bar{A}^{O-1}) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * ((A * L) \otimes (A * L)) + (I_{p \times p} \otimes (C * \bar{A}^{O-2} * A * L)) \end{bmatrix}. \quad (5.29)$$

V rovnici (5.26) je také ještě vektor neznámého parametru $\theta(Q, R)$, který obsahuje všechny prvky matic Q a R . Dále je v rovnici (5.26) ještě vektor pravé strany b , který je tvořen kovariancemi inovační posloupnosti a má následující podobu:

$$b = \begin{bmatrix} E[z_k * z_k^T] \\ E[z_k * z_{k+1}^T] \\ \vdots \\ E[z_k * z_{k+O-1}^T] \end{bmatrix}_S. \quad (5.30)$$

Celkový počet kovariančních rovnic je roven O . Všechny tyto kovariance ale neznáme. Můžeme však vypočítat jejich odhad pomocí rovnice (2.32). Při znalosti nulové střední hodnoty inovační posloupnosti lze vypočítat nestranný odhad například první kovariance dle následujícího vztahu:

$$E[\widehat{z_k * z_k^T}] = \frac{1}{N-1} * \sum_1^N z_k * z_k^T. \quad (5.31)$$

Nebo pro libovolný časový posun j inovační posloupnosti lze vypočítat odhad kovariance stejně jako v rovnici (2.32) dle následujícího vztahu:

$$E[\widehat{z_k * z^T_{k+j}}] = \frac{1}{N-1-j} * \sum_1^N z_k * z^T_{k+j}. \quad (5.32)$$

Těmito výpočty (5.31) a (5.32) lze vypočítat odhad celé pravé strany rovnice (5.26), která bude tedy vypadat následovně:

$$\hat{b} = \begin{bmatrix} E[\widehat{z_k * z^T_k}] \\ E[\widehat{z_k * z^T_{k+1}}] \\ \vdots \\ E[\widehat{z_k * z^T_{k+O-1}}] \end{bmatrix}_S. \quad (5.33)$$

Nyní stačí pouze za vektor b dosadit nestranný odhad \hat{b} a vyřešit soustavu lineárních rovnic (5.26) pomocí metody nejmenších čtverců. Díky metodě nejmenších čtverců (viz Příloha) lze ukázat, že pokud do rovnice (5.26) dosadíme odhad \hat{b} , výsledný odhad neznámého parametru $\theta(Q, R)$ je nestranný.

5.2 Vlastnosti metody

Jednou z velkých výhod této metody je jistě fakt, jak už bylo řečeno, že metoda je analyticky odvozená a dává nestranné odhady kovariančních matic.

Dále pak tu však máme i vlastnosti, které působí určité problémy v odhadu či jejich výsledné kvalitě. Vlastnost, která ovlivňuje kvalitu odhadu, je jistě volba konstantního zisku L v rovnici (5.9). Volbu L ale nelze dělat libovolně. Musíme zaručit stabilitu stavové rovnice chybového modelu (5.12). Na ukázkou bude uveden příklad, který bude prezentovat volbu zisku L na kvalitu odhadu. Zvolíme tedy systém:

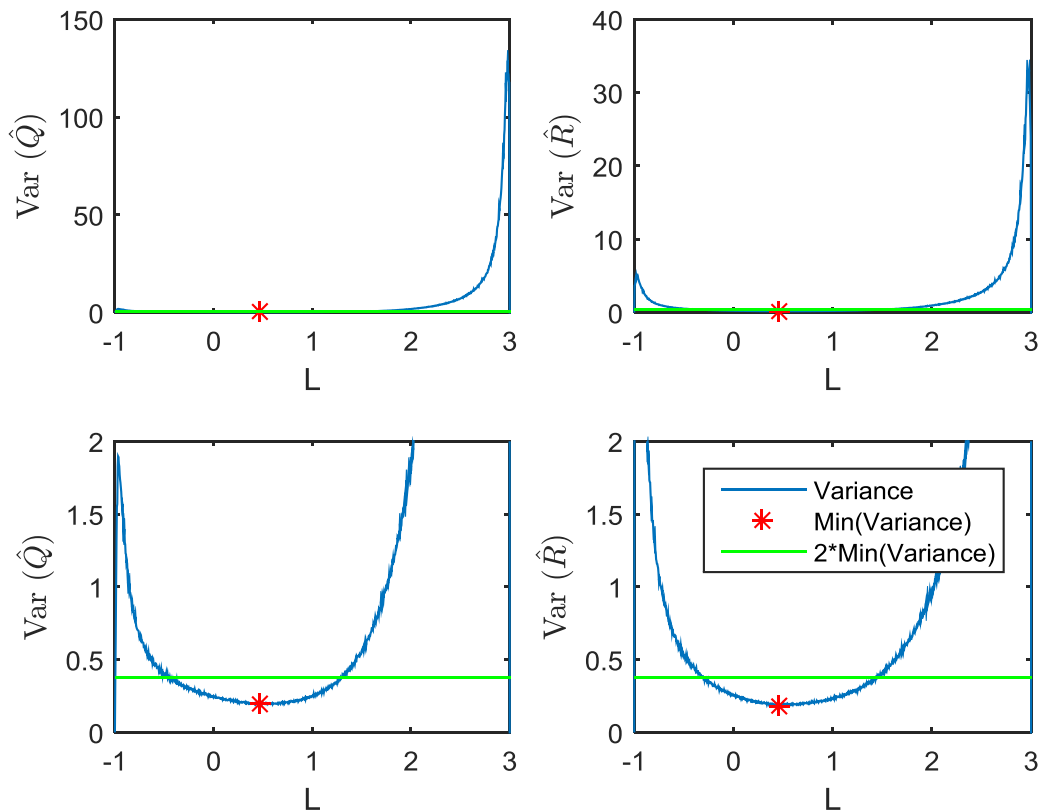
$$A = 0.5; C = 1; Q = 1; R = 1. \quad (5.34)$$

Pro takto zvolený systém můžeme zisk L volit v rozmezí od -1 do 3, aby stavová rovnice chybového modelu (5.12) byla stále stabilní. To se dá jednoduše ukázat následujícím výpočtem:

$$\begin{aligned} |A - A * L * C| &< 1 \\ |0.5(1 - L)| &< 1 \\ -1 &< L < 3. \end{aligned} \quad (5.35)$$

Je také důležité říci, že tento výpočet lze provést pouze pro systém prvního řádu.

Pokud tedy nasimulujeme metodu ACLS pro 100 kroků systému, 10000 Monte Carlo simulací a $O = 4$ kovarianční rovnice, pro takto zvolený systém bude kvalita odhadu, která byla volena jako variance odhadu, vypadat následovně:



Obrázek 6 – Variance odhadu v závislosti na volbě L

kde na prvních dvou grafech v obrázku 6 je vidět variance odhadu matic Q a R , je zde také vykreslena minimální variance a pro představu také, jak se zhorší odhad při špatné volbě zisku i dvojnásobek minimální variance. Na dalších dvou grafech je pouze přiblížení prvních dvou grafů. Je tedy vidět, že volba zisku ovlivní výslednou kvalitu odhadu. Je také důležité si všimnout, že minimum variance odhadu matice Q a matice R jsou na stejném místě, proto zde existuje pouze jeden zisk L , pro který bude variance odhadu pro obě kovarianční matice minimální.

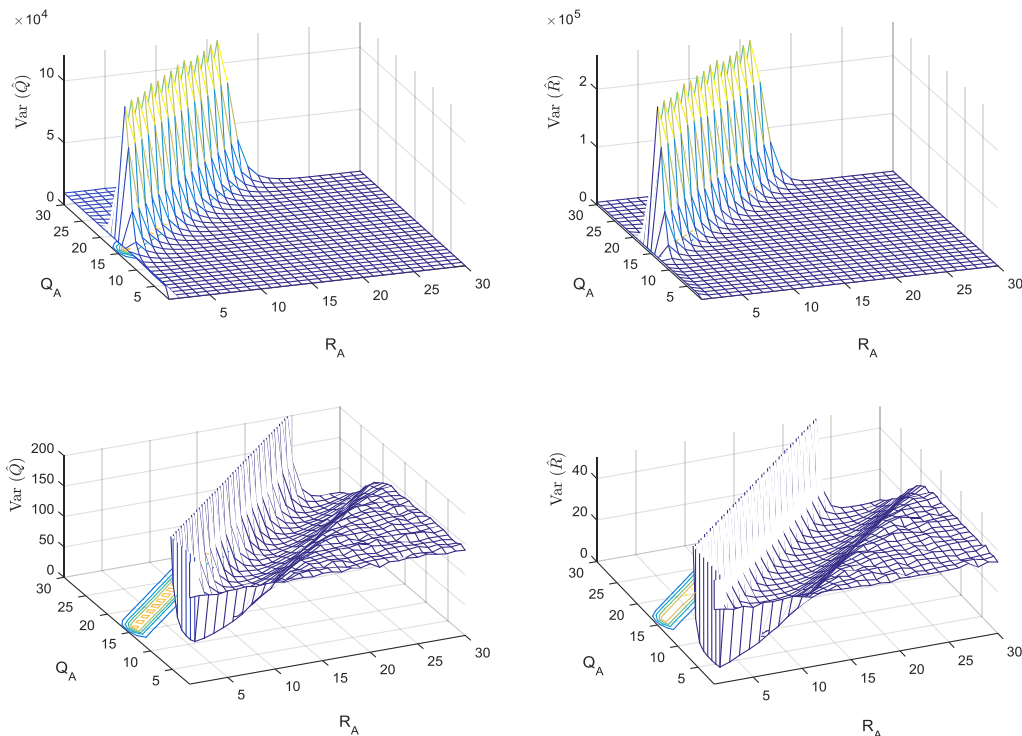
Pokud máme systém prvního řádu, je možné určit omezení volby zisku jako v rovnici (5.35) a lze tedy volit zisk „pouze“ z takto zvoleného intervalu. Pokud bychom měli systém vyššího řádu, bude volba zisku obtížnější.

Pro příklad bude ukázán odhad vlastností poruch pro systém, který je dán následujícími maticemi:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.324 & 0.81 & -0.4 \end{bmatrix}; C = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 1 & 0.5 \\ 1.2 & 0.4 & 0.1 \end{bmatrix}; \quad (5.36)$$

$$Q = \begin{bmatrix} 1 & 0.5 & 0.6 \\ 0.5 & 3 & -0.3 \\ 0.6 & -0.3 & 2 \end{bmatrix}; R = \begin{bmatrix} 0.8 & 0.4 & 0.3 \\ 0.4 & 1 & 0.1 \\ 0.3 & 0.1 & 2 \end{bmatrix}.$$

Zisk L byl zde vypočítáván dle rovnice (3.14) a (3.15), kde za matice Q a R , které bereme jako neznámé, byly postupně dosazovány matice Q_A a R_A . Pro příklad matice Q_A se počítala pomocí vztahu $Q_A = I_{n \times n} * c$, kde $c \in \{0, \{2^{i-1}\}_{i=1}^{30}\}$. Stejně tak se počítala i matice R_A , jen měla jednotkovou matici dimenze $p \times p$. Simulace systému bude opět prováděna pro $N = 100$ kroků, 10000 Monte Carlo simulací a $O = 4$ kovariančních rovnice. Výsledné variance odhadů vypadají následovně:

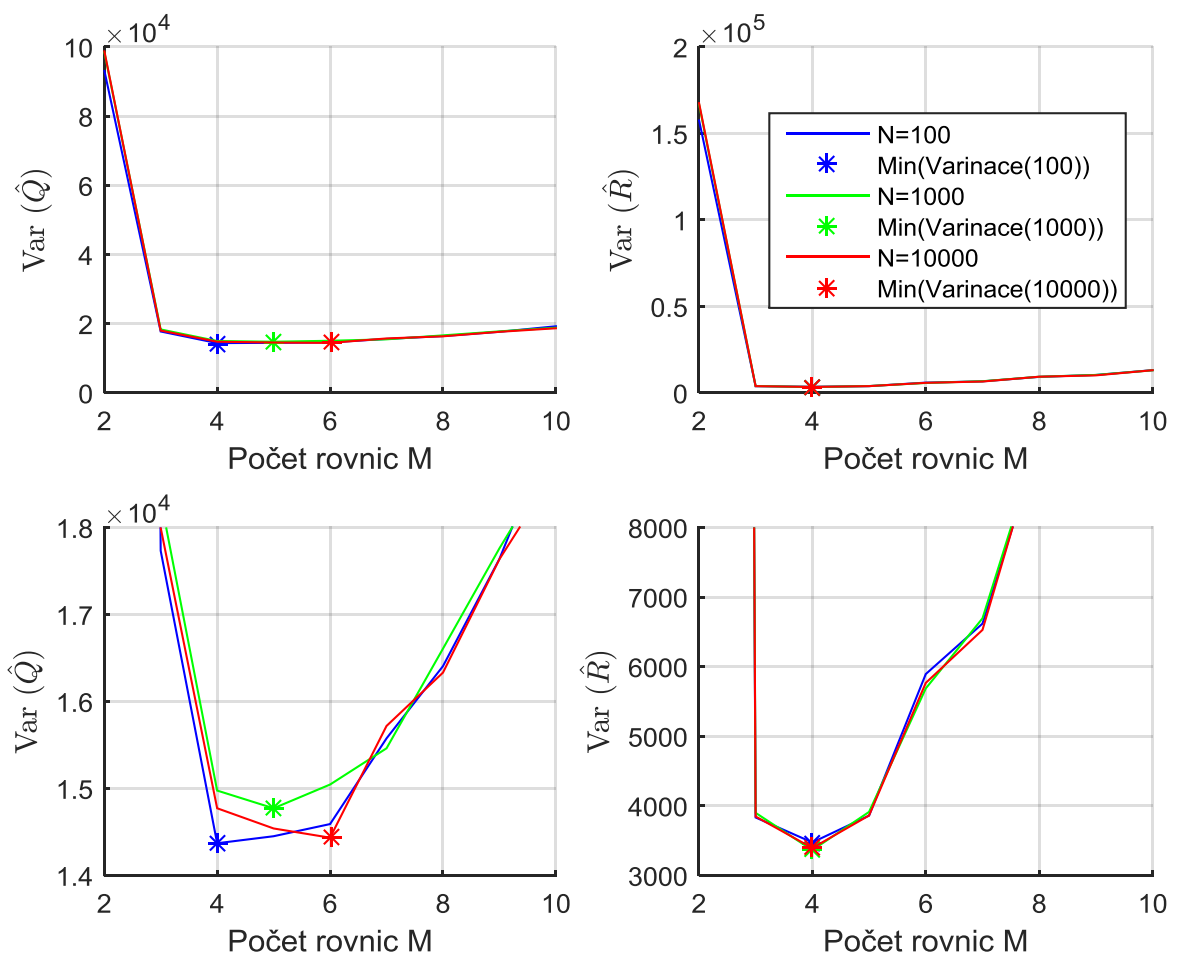


Obrázek 7 - Variance odhadu v závislosti na volbě Q_A a R_A

kde na prvních dvou grafech v obrázku 7 je vidět součet unikátních prvků variancí odhadů matic Q a R . Na dalších dvou grafech je pouze přiblížení prvních dvou grafů, kde na ose Q_A a R_A jsou hodnoty indexu i , jež byl využit pro jejich výpočet. Zde je vidět stejně jako v předchozím případě, že variance odhadu opět závisí na volbě zisku

L . Je také vidět, že pokud zisk počítáme z rovnice (3.14) a (3.15) jako v tomto případě, nezáleží na velikostech matic Q_A a R_A , ale zejména na jejich poměru. Otázkou tedy zůstává, jak nalézt optimální zisk L tak, aby výsledná variance odhadu byla minimální.

Další vlastností, která působí určité problémy v odhadu, je volba počtu kovariančních rovnic O . Jako příklad bude opět simulován systém popsáný v rovnici (5.36), kde bude zisk volený pro jednoduchost nulový a počet Monte Carlo simulací bude 10000. Výsledky simulace vypadají následovně:



Obrázek 8 - Variance odhadu v závislosti na volbě N a počtu kovariančních rovnic

kde na prvních dvou grafech v obrázku 8 je vidět variance odhadu matic Q a R při různých počtech rovnic a kroků N . Na dalších dvou grafech je pouze přiblížení prvních dvou grafů. Jednotlivé variance jsou vynásobeny počtem kroků N , to je proto, aby bylo dobře vidět, že pro systém, který budeme měřit desetkrát déle, pomocí metody ACLS dostaneme přibližně desetkrát menší varianci odhadu. To znamená, že pokud chceme identifikovat kovarianční matice poruch, měli bychom počet kroků N volit co největší. Z obrázku 8 je také patrné, že pro nejmenší varianci odhadu matice R

je nejlepší volit čtyři rovnice, kdežto u matice Q vycházejí nejmenší variance odhadu v průměru pro pět rovnic. Proto byly čtyři rovnice určeny jako vhodná volba. Tyto závěry vycházely i z jiných simulací pro jiné systémy než (5.36).

Poslední vlastností nebo spíše omezením je maximální počet identifikovatelných prvků. Metoda ACLS odhadne všechny prvky matic Q a R za předpokladu, že budeme mít dostatečný počet naměřených dat, ze kterých se utváří (5.33), a pouze pokud matice \mathcal{A} v rovnici (5.26) má plnou hodnotu. Jak už bylo psáno v rovnici (5.27), matici lze rozdělit na dvě části \mathcal{A}_1 a \mathcal{A}_2 , z nichž \mathcal{A}_1 ovlivňuje odhad matice Q a \mathcal{A}_2 odhad matice R .

Je důležité upozornit na následující vztahy, které nám definují pravidla pro výpočet hodnosti matic:

$$\text{rank}([S + D]) \leq \text{rank}(S) + \text{rank}(D), \quad (5.37)$$

$$\text{rank}(S * D) \leq \min[\text{rank}(S); \text{rank}(D)], \quad (5.38)$$

$$\text{rank}(S \otimes D) \leq \text{rank}(S) * \text{rank}(D), \quad (5.39)$$

kde S a D jsou libovolné matice.

Nejdříve se zaměříme na hodnotu matice \mathcal{A}_1 . Ta je díky rovnici (5.38) dána vztahem

$$\text{rank}(\mathcal{A}_1) \leq \min[\text{rank}(C \otimes C * \bar{A}^j); \text{rank}((I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1})]. \quad (5.40)$$

Pokud je systém (2.33), (2.34) pozorovatelný, což je také jednou z podmínek použití jak už Kalmanova filtru, tak i samotné metody ACLS, a matice \bar{A} stabilní, lze říci, že matice $(C \otimes C * \bar{A}^j)$ bude mít, jak je vidět z rovnice (5.39), maximálně hodnotu $n * p$. Dále matice $(I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1}$ má za předpokladu stabilní matice \bar{A} hodnotu n^2 . Což znamená, že matice \mathcal{A}_1 dle rovnice (5.40) může mít maximální hodnotu $n * p$, z čehož vyplývá, že dokážeme odhadnout $n * p$ prvků matice Q , jež má n^2 prvků. Pokud tedy budeme chtít odhadnout všechny prvky matice Q , musíme mít více nebo stejně rovnic měření, než máme řád systému.

Dále se podíváme na hodnotu matice \mathcal{A}_2 . Ta je opět díky rovnici (5.38) a (5.37) dána následujícím vztahem:

$$\text{rank}(\mathcal{A}_2) \leq \min \left[\begin{array}{l} \text{rank}(C \otimes (C * \bar{A}^j)) \\ \text{rank}((I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1}) \\ \text{rank}((A * L) \otimes (A * L)) \end{array} \right] + \text{rank}(I_{p \times p} \otimes C * \bar{A}^{j-1} * A * L). \quad (5.41)$$

Jak už bylo řečeno, matice $(C \otimes (C * \bar{A}^j))$ má hodnotu $n * p$ a matice $(I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1}$ má hodnotu n^2 . Matice $((A * L) \otimes (A * L))$ se skládá ze dvou matic, které při vynásobení mohou mít maximální hodnotu rovnou p^2 . První člen rovnice (5.41), ve které jsme hledali minimální hodnotu, je tedy roven p^2 . Druhá matice $(I_{p \times p} \otimes (C * \bar{A}^{j-1} * A * L))$ má za předpokladu opětovného předpokladu „libovolné“ volby L hodnotu p^2 . Což tedy znamená, že výsledná hodnota matice $\text{rank}(\mathcal{A}_2)$ je z rovnice (5.37) maximálně $2 * p^2$. To znamená, že můžeme odhadnout všechny prvky matice R , kterých je p^2 .

Z předchozích výpočtů je vidět, že v matici R se dají odhadnout všechny prvky, ale u matice Q , pokud máme méně rovnic měření než systému, pouze $n * p$. Proto bylo provedena simulace systému třetího řádu se dvěma měřeními. Matice Q má devět prvků, ale lze odhadnout pouze šest, to se potvrdilo a odhad celé matice Q byl nesprávný. Proto bylo využito poznatku, že matice Q je symetrická a má pouze šest unikátních prvků, a proto byly sjednoceny vektory matice \mathcal{A}_1 , jež odpovídaly stejným křížovým prvkům. To znamená, že výsledná velikost matice \mathcal{A}_1 neměla devět sloupců, ale šest, což odpovídá maximálnímu počtu možných odhadnutelných prvků matice Q . Ukázalo se, že minimalizací matice \mathcal{A}_1 na unikátní prvky se však snížila hodnota matice z původních šesti na pět. To znamená, že v tomto případě celkem dokážeme odhadnout maximálně osm prvků, pokud jako neodhadovaný prvek byl prvek na diagonále, a sedm prvků, pokud by to byl prvek mimo diagonálu. Neodhadovaný prvek by se ovšem musel zvolit, což pokud bychom tento prvek zvolili tak, že by neodpovídal skutečné hodnotě prvku, nebyl by výsledný odhad nestranný.

6. Kovarianční matice odhadu a její horní mez

Tato kapitola se zabývá tím, jak zvolit zisk L tak, aby metoda ACLS dávala odhady s minimální variancí.

6.1 Odvození metody

Na začátku předchozí kapitoly 5.2 byl na obrázku 6 a 7 ukázán problém volby lineárního zisku L a to, jak tato volba ovlivňuje kvalitu odhadu. Pokud jsme zvolili jako ukazatel kvality odhadu jeho varianci, bude vhodné pro volbu zisku L nějakým způsobem vypočítat odhad této variance. Nejdříve se tedy podíváme, jak vypadá variance odhadu metody ACLS, která se vypočte následovně:

$$\begin{aligned} cov[\hat{\theta}] &= E[(\theta - \hat{\theta}) * (\theta - \hat{\theta})^T] = \\ &= E[(\mathcal{A}^\dagger * b - \mathcal{A}^\dagger * \hat{b}) * (\mathcal{A}^\dagger * b - \mathcal{A}^\dagger * \hat{b})^T] = \quad (6.1) \\ &= \mathcal{A}^\dagger * E[(b - \hat{b}) * (b - \hat{b})^T] * \mathcal{A}^{\dagger T} = \mathcal{A}^\dagger * cov[\hat{b}] * \mathcal{A}^{\dagger T}, \end{aligned}$$

kde matice \mathcal{A}^\dagger značí pseudoinverzi matice \mathcal{A} (viz Appendix). Matice \mathcal{A} popsaná v rovnici (5.27) je známá. Problém je tedy se znalostí kovariance vektoru \hat{b} popsané v rovnici (5.33). Výpočet kovarianční matice odhadu (6.1) nelze vypočítat, neboť neznáme matice Q a R ani optimální zisk L . Co však lze vypočítat, je horní mez omezující kovarianční matici odhadu. Ta pak může být minimalizována vzhledem k L . Po této minimalizaci věříme, že argument tohoto minima bude odpovídat optimálnímu zisku L .

Nejprve se podívejme na kovarianci \hat{b} pouze pro první kovarianční rovnici, při které nebyl žádný časový posuv j . Pro zjednodušení bude první rovnice značená následovně:

$$E[\widehat{z_k * z_k^T}]_S = \hat{C}_S. \quad (6.2)$$

Kovariance rovnice (6.2) se dá vypočítat pomocí vzorce (2.27), jak je vidět v následujícím vztahu:

$$cov[\hat{C}_S] = E[\hat{C}_S * \hat{C}_S^T] - E[\hat{C}_S] * E[\hat{C}_S]^T. \quad (6.3)$$

Nejdříve se v rovnici (6.3) zaměříme na první část výpočtu, a to $E[\hat{C}_S * \hat{C}_S^T]$, kterou pomocí rovnice (2.32) vypočítáme následovně:

$$\begin{aligned} E[\hat{C}_S * \hat{C}_S^T] &= \frac{1}{N^2} * E\left[\left(\sum_{i=1}^N z_i * z_i^T\right)_S * \left(\sum_{j=1}^N z_j * z_j^T\right)_S^T\right] = \\ &= \frac{1}{N^2} * E\left[\sum_{i=1}^N \sum_{j=1}^N (z_i * z_i^T)_S * (z_j * z_j^T)_S^T\right]. \end{aligned} \quad (6.4)$$

Pro zjednodušení zavedeme následující substituci:

$$\alpha_i = (z_i * z_i^T)_S. \quad (6.5)$$

Poté se dá rovnice (6.4) přepsat a upravit do následující podoby:

$$E[\hat{C}_S * \hat{C}_S^T] = \frac{1}{N^2} * \left[\sum_{i=1}^N E[\alpha_i * \alpha_i^T] + \sum_{i=1}^N \sum_{\substack{j=1 \\ i \neq j}}^N E[\alpha_i * \alpha_j^T] \right]. \quad (6.6)$$

Je důležité si také uvědomit, že kovariance jednoho prvku sama se sebou ve stejném okamžiku bude mít vždy větší hodnotu než kovariance dvou prvků nebo jednoho v různých časových okamžicích. Kovarianci dvou prvků nebo jednoho v různých časových okamžicích se též říká křížová variance. Poznatek o velikosti kovariance jednoho prvku ve stejném okamžiku a křížové kovarianci je uveden v následujícím vztahu:

$$cov[\alpha_i] = cov[\alpha_i; \alpha_i] \geq cov[\alpha_i; \alpha_j], \forall i, j. \quad (6.7)$$

Dále je také důležité upozornit že inovační posloupnost je gaussovská. To znamená, že se dá plně popsat prvními dvěma momenty a další lze dopočítat ze znalosti těchto prvních dvou momentů. Čtvrtý moment proto lze vypočítat následovně:

$$E[(z_i * z_i^T) * (z_i * z_i^T)^T] = E[\alpha_i * \alpha_i^T] \cong 3 * C_S * C_S. \quad (6.8)$$

Výsledek rovnice (6.8) exaktně platí pro systém prvního řádu. Pro systémy vyšších řádů se jedná pouze o aproximaci.

Pokud bychom tedy poznatek z rovnice (6.7) využili v rovnici (6.6), nahradili bychom křížové prvky nekřížovými a využili bychom také rovnici (6.8), rovnice (6.6) by vypadala následovně:

$$\text{cov}[\hat{C}_S] \leq \frac{1}{N^2} [N^2 * 3 * C_S * C_S] - E[\hat{C}_S] * E[\hat{C}_S]^T, \quad (6.9)$$

kde vektor \hat{C}_S je nestranný odhad C_S a tím pádem platí, že $E[\hat{C}_S]$ se rovná C_S . Tudiž z rovnice (6.9) lze vyvodit vztah:

$$\text{cov}[\hat{C}_S] \leq \text{cov}[\hat{C}_S]_{up} \leq 2 * C_S * C_S, \quad (6.10)$$

kde index up značí, že se jedná o horní mez kovariance vektoru \hat{C}_S díky aproximaci rovnice (6.7).

Z rovnice (5.23) využijeme hned počáteční vztah:

$$C_S = (C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * (\bar{G} \otimes \bar{G}) * \bar{Q}_S + R_S. \quad (6.11)$$

V této rovnici ovšem neznáme vektory R_S a \bar{Q}_S . Proto zavedeme následující vztah:

$$M * \Sigma_S = [(C \otimes C) * (I_{n^2 \times n^2} - \bar{A} \otimes \bar{A})^{-1} * (\bar{G} \otimes \bar{G}) + J_t] * \Sigma_S, \quad (6.12)$$

kde vektor Σ_S reprezentuje neznámý vektor \bar{Q}_S obsahující jak vektor Q_S , tak i R_S . Matice J_t má specifické rozložení nul a jedniček takové, aby při vynásobení vektorem Σ_S byly vybrány pouze prvky vektoru R_S . Kupříkladu pokud má systém dvě měření ($p = 2$), matice J_t vypadá následovně:

$$J_t = \begin{bmatrix} 0_{p^2 \times n^2+n*p+n} & I_{p \times p} & 0_{p^2 \times n} & 0_{p \times p} \\ 0_{p \times p} & 0_{p \times p} & 0_{p^2 \times n} & I_{p \times p} \end{bmatrix}, \quad (6.13)$$

nebo tři měření ($p = 3$), matice J_t vypadá následovně:

$$J_t = \begin{bmatrix} 0_{p^2 \times n^2+n*p+n} & I_{p \times p} & 0_{p^2 \times n} & 0_{p \times p} \\ 0_{p^2-p \times p} & 0_{p^2-p \times p} & 0_{p^2 \times n} & I_{p \times p} \\ 0_{p \times p} & 0_{p \times p} & 0_{p^2 \times n} & 0_{p^2-p \times p} \\ 0_{p \times p} & 0_{p \times p} & 0_{p^2 \times n} & I_{p \times p} \end{bmatrix}, \quad (6.14)$$

kde matice 0 je matice plná nul s dimenzí totožnou s jejím spodním indexem.

Rovnice (6.12) nám tedy nabízí jakousi náhradu za kovarianci C_S . Toho můžeme využít spolu s předchozím poznatkem, že velikosti kovariance jednoho prvku ve stejném okamžiku a křížové kovariance lze tedy zavést následujícím vztahem:

$$\begin{aligned} cov[\hat{b}]_{up} &= P_c * (J_n \otimes cov[\hat{C}_S]_{up}) * P_c^T = \\ &= P_c * (2 * (J_n \otimes (M * \Sigma_S * \Sigma_S^T * M^T))) * P_c^T, \end{aligned} \quad (6.15)$$

kde J_n je matice jedniček takové dimenze, jako je počet kovariančních rovnic, tj. O , a P_c je permutační matice jedniček a nul, která správně přeuspořádá prvky výsledné horní meze kovarianční matice.

Nyní tedy máme maximální kovarianci odhadu vektoru pravé strany. Máme tu ale neznámý vektor Σ_S , který je tvořen maticemi Q a R . Matice $\Sigma_S * \Sigma_S^T$ bude pozitivně definitní. Proto a také kvůli tomu, že tyto matice neznáme, je vynecháme a vzorec (6.15) redukuje na následující tvar:

$$J(L) = \mathcal{A}^\dagger * P_c * (J_n \otimes (M * M^T)) * P_c^T * \mathcal{A}^{\dagger T}. \quad (6.16)$$

Pokud tuto funkci představující horní mez kovarianční matice odhadu minimalizuje vzhledem k L , věříme, že takto nalezený zisk použitý v metodě ACLS nám umožní získat velice kvalitní odhad ve smyslu jeho variance. Zisk se budeme snažit najít pomocí následujícího vzorce:

$$L = \underset{L}{argmin}(J(L)). \quad (6.17)$$

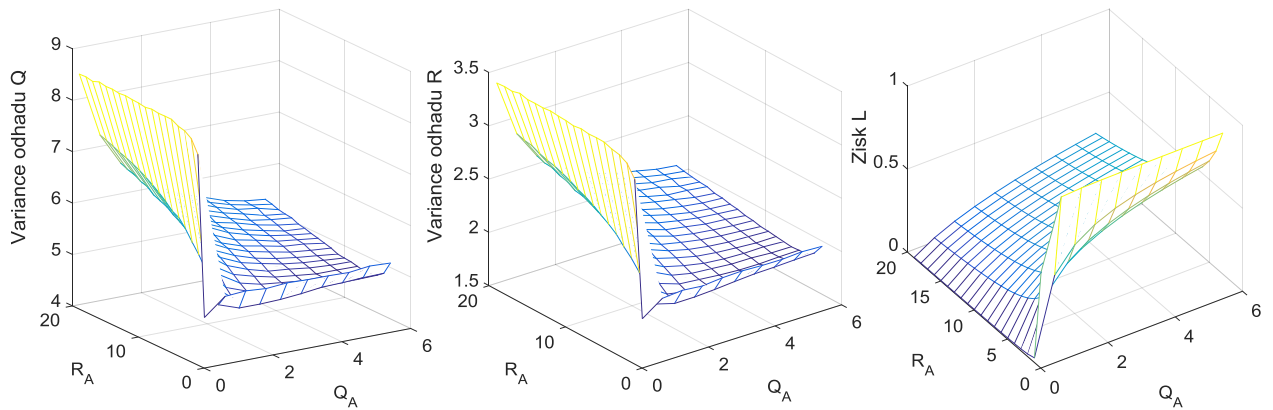
6.2 Numerická ilustrace

Pro ověření kritéria, jež bylo navrženo v předchozí kapitole, zde bude ukázáno několik příkladů, které ukážou, zda je volba zisku nalezená pomocí kritéria vhodná.

Jako první zde uvedeme příklad pro tento systém:

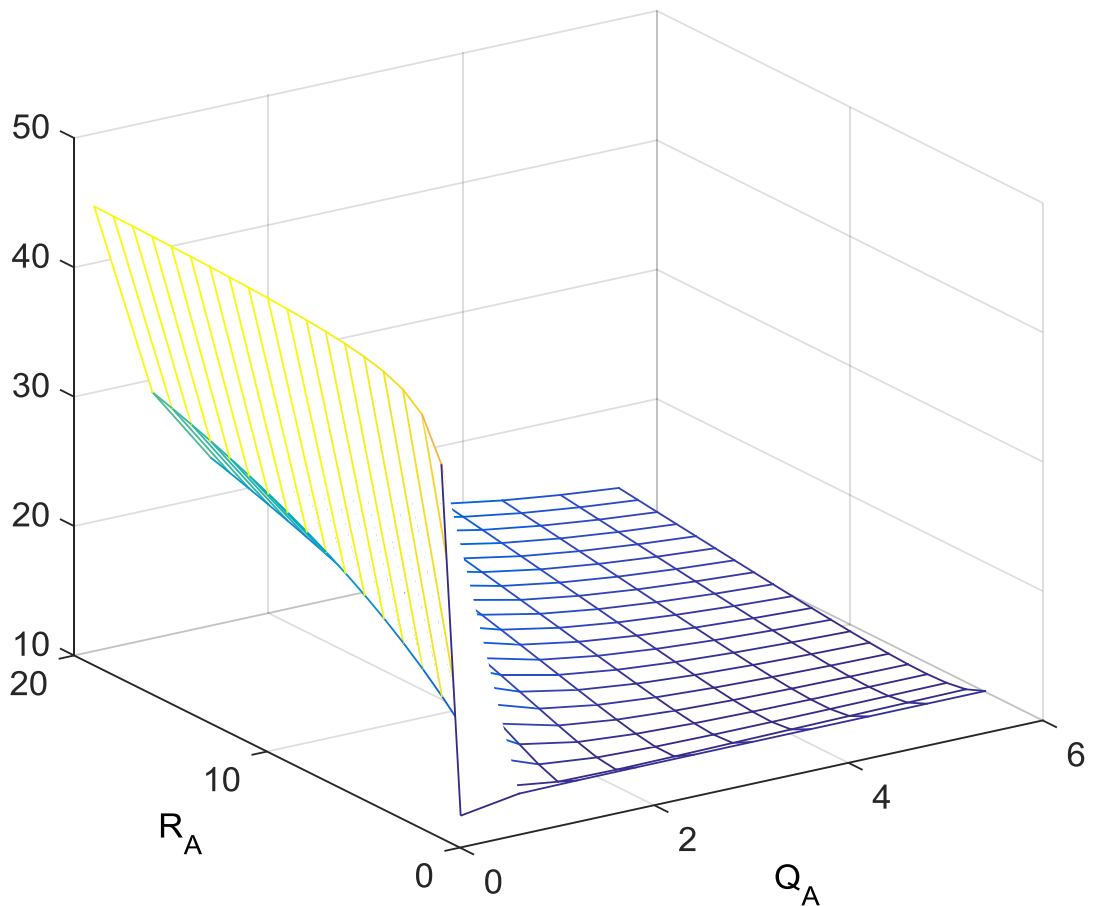
$$A = -0.8; C = 1; Q = 8; R = 2. \quad (6.18)$$

Tento systém bude prováděn pro $N = 100$ kroků, 10000 Monte Carlo simulací a $O = 4$ kovariančních matic. Výsledné variance odhadů vypadají následovně:



Obrázek 9 - Variance odhadů Q a R a zisk L v závislosti na volbě Q_A a R_A

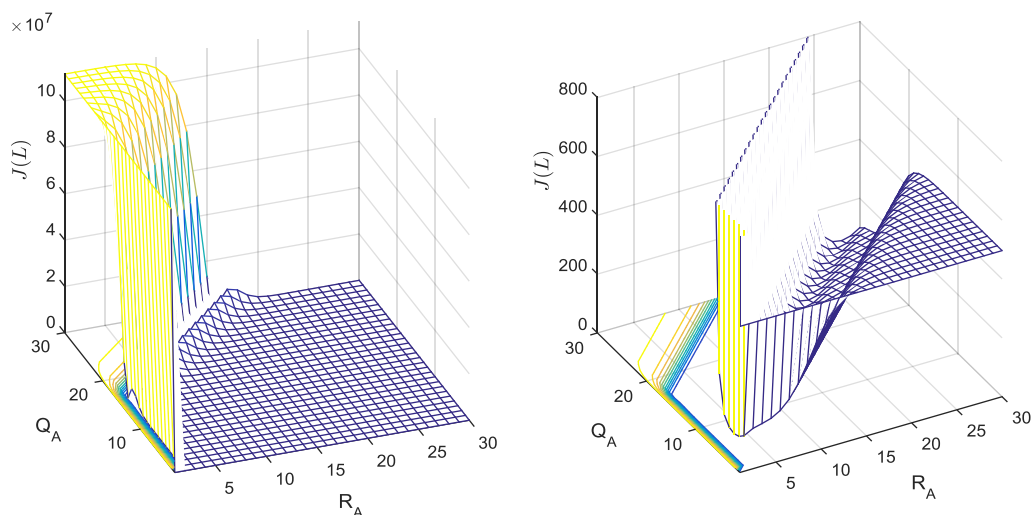
kde zisk L zde byl vypočítáván dle rovnice (3.14) a (3.15), kde za matice Q a R , které bereme jako neznámé, byly postupně dosazovány matice $Q_A \in \{0.01, 0.5, 1, \dots, 6\}$ a $R_A \in \{0.01, 0.5, 1, \dots, 20\}$. Výsledný výpočet dle kritéria (6.16) vypadá následovně:



Obrázek 10 - Hodnota kritéria J v závislosti na volbě Q_A a R_A , $n=1$

Na obrázku 10 je kritérium vykresleno jako součet prvků na diagonále matice $J(L)$. Z obrázku 10 je vidět, že tvar přibližně odpovídá skutečným variancím odhadu prvků Q a R v obrázku 9. Proto se zdá, že toto kritérium by pro tento systém bylo vyhovující a nalezený optimální zisk z rovnice (6.17) by byl vhodný i pro získání kvalitních odhadů v metodě ACLS.

Jako druhý příklad použijeme vícedimenzionální systém navržený v rovnici (6.36):



Obrázek 11 - Hodnota kritéria J v závislosti na volbě Q_A a R_A , $n=3$

kde na obrázku 11 je kritérium vykresleno jako součet prvků na diagonále matice $J(L)$ a kde na ose Q_A a R_A jsou hodnoty indexu i , jež byl využit pro jejich výpočet. Z obrázku 11 je stejně jako na obrázku 10 vidět, že tvar přibližně odpovídá skutečným variancím odhadu prvků Q a R . I v tomto případě je kritérium pro tento systém vyhovující a nalezený optimální zisk z rovnice (6.17) by byl opět vhodný pro získání kvalitních odhadů pomocí metody ACLS. Z obrázku 11 je také vidět, že v části, kde bylo voleno $Q_A = I_{n \times n} * 2^{29}$ a $R_A = 0_{p \times p}$, kritérium neodpovídá obrázku 7. Je to nejspíše zapříčiněno aproximací v rovnici (6.8). Proto volba zisku L za pomoci kritéria v rovnici (6.17) je, spíše než platným pravidlem, jakýmsi doporučením jak tento zisk volit. Nelze totiž zaručit, že nesrovnalost zapříčiněná aproximací se neprojeví právě v minimu kritéria $J(L)$, a neovlivní tak i optimální volbu zisku L pro metodu ACLS. Na závěr je také důležité říci, že pro mnoho simulací a testování této metody je jakousi vhodnou volbou nulový zisk L . Při všech možných testováních pro různé systémy se ukázalo, že tato volba není optimální, ale vždy dávala uspokojivé výsledky – viz obrázky 7 a 9.

7. Odhad časově korelovaných poruch

Tato kapitola se zabývá rozšířenou metodou ACLS, jejím odvozením, kontrolou správného odhadu a porovnáním s odhadem vypočteným pomocí metody založené na Allanově varianci.

7.1 Odvození metody

Metoda ACLS byla v původním znění v [3] navržena pro systém s bílými poruchami tak, jak je uvažováno v (2.33) a (2.34). V této kapitole bude ukázána možnost tento předpoklad rozšířit o složitější poruchy působící na rovnici měření, ve smyslu jejich struktury jako v [23] – [28]. Bělost poruchy v rovnici (2.33) zůstane nezměněna, ale rovnice měření (2.34) se změní na následující:

$$y_k = C * x_k + v_k + u_k + q, \quad (7.1)$$

kde $u \in \mathbb{R}^{p \times 1}$ je bílý šum popsáný jako $N(0, R_u)$, $q \in \mathbb{R}^{p \times 1}$ je konstantní přírůstek neboli bias a $v \in \mathbb{R}^{p \times 1}$ je časově korelovaný šum neboli Gauss-Markovský proces popsáný $N(0, R_v)$ s následující vnitřní strukturou:

$$v_{k+1} = \lambda * v_k + \xi_k, \quad (7.2)$$

kde $\xi \in \mathbb{R}^{p \times 1}$ je bílý šum popsáný jako $N(0, R_\xi)$ a stabilní parametr $\lambda \in \mathbb{R}^{p \times p}$.

Jako příklad senzoru, který je ovlivněn poruchou tohoto typu, lze zmínit například inerciální senzory, jako je akcelerometr či gyroskop. Tato rozšířená metoda ACLS využívá možnosti zvolit zisk L rovný nule, což znamená, že matice systému A musí být stabilní. Pokud totiž zisk zvolíme nulový, veškeré vztahy se velice zjednoduší a rovnice (5.1) se změní na:

$$\hat{x}_{k+1|k} = A * \hat{x}_{k|k-1} + B * u_k. \quad (7.3)$$

Pokud bychom tuto rovnici dále dosadili do zavedeného vztahu (5.4) s využitím rovnice systému (2.33), získáme následující rovnici:

$$\begin{aligned}\varepsilon_{k+1} &= A * x_k + B * u_k + w_k - A * \hat{x}_{k|k-1} - B * u_k \\ \varepsilon_{k+1} &= A * (x_k - \hat{x}_{k|k-1}) + w_k\end{aligned}\quad (7.4)$$

$$\varepsilon_{k+1} = A * \varepsilon_k + w_k.$$

Dále si ukážeme, že chyba odhadu stavu ε má stále střední hodnotu nulovou následujícím výpočtem:

$$E[\varepsilon_{k+1}] = E[A * \varepsilon_k + w_k] = A * E[\varepsilon_k] = 0, \quad (7.5)$$

neboť střední hodnota poruch $E[\bar{w}_k]$ je rovna nule a za předpokladu stability matice A se střední hodnota chyby prediktivního odhadu stavu ustálí, tudíž bude platit $E[\varepsilon_{k+1}] = E[\varepsilon_k]$, tím pádem se střední hodnota chyby prediktivního odhadu bude rovnat nule.

Jako další se nám bude hodit kovariance ε , kterou vypočítáme následujícím výpočtem:

$$\begin{aligned}cov[\varepsilon_{k+1}] &= E[(A * \varepsilon_k + w_k) * (A * \varepsilon_k + w_k)^T] = \\ &= E[A * \varepsilon_k * \varepsilon_k^T * A^T + w_k * w_k^T] = A * cov[\varepsilon_k] * A^T + Q.\end{aligned}\quad (7.6)$$

Pokud bude matice A stabilní, bude mít tato Ljapunovova rovnice ustálené řešení P^- . Toto ustálené řešení se vypočte následujícím výpočtem:

$$\begin{aligned}P^- &= A * P^- * A^T + Q, \\ P^-_S &= (I_{n^2 \times n^2} - A \otimes A)^{-1} Q_S.\end{aligned}\quad (7.7)$$

Z rovnice (7.4) jsme tedy získali rovnici systému a stejně jako v předchozím případě musíme nadefinovat i výstupní rovnici, která vypadá následovně:

$$\begin{aligned}z_k &= C * \varepsilon_k + v_k + u_k + q = C * x_k + v_k + u_k + q - C * \hat{x}_{k|k-1} = \\ &= y_k - \hat{y}_{k|k-1} + q.\end{aligned}\quad (7.8)$$

Je zřejmé, že oproti výstupní rovnici (5.15) se v rovnici objevil bias q . Ten ovšem definuje střední hodnotu výstupní rovnice (7.8), jak je vidět v následujícím výpočtu:

$$\begin{aligned}
E[z_k] &= E[y_k - \hat{y}_{k|k-1} + q] = \\
&= E[C * x_k + v_k + u_k + q - C * \hat{x}_{k|k-1} + \hat{q} + q] = \\
&= E[C * x_k + q - C * \hat{x}_{k|k-1} - \hat{q} + q] = q.
\end{aligned} \tag{7.9}$$

Jak bude později ukázáno, rovnice (7.9) je klíčová pro odhad parametru q . Díky znalosti tohoto parametru lze tedy nadefinovat alternativní rovnici výstupu, která vypadá následovně:

$$z_k^* = z_k - E[z_k] = z_k - q = C * \varepsilon_k + v_k + u_k = y_k - \hat{y}_{k|k-1}. \tag{7.10}$$

Takto definovaný výstup již odpovídá výstupu definovaného v rovnici (5.15).

Dále stejně jako u metody ACLS musíme vypočítat kovariance inovační posloupnosti neboli také výstupní rovnice (7.10). První kovariance výstupu ve stejném časovém okamžiku lze tedy vypočítat následovně:

$$\begin{aligned}
E[z_k^* * z_k^{*T}] &= E[(C * \varepsilon_k + v_k + u_k) * (C * \varepsilon_k + v_k + u_k)^T] = \\
&= E[C * \varepsilon_k * \varepsilon_k^T * C^T + v_k * v_k^T + u_k * u_k^T] = \\
&= C * P^- * C^T + R_v + R_u.
\end{aligned} \tag{7.11}$$

Další kovariance výstupu (7.10), nyní pro okamžiky posunuté o jeden časový index, vypadá následovně:

$$\begin{aligned}
E[z_k^* * z_{k+1}^{*T}] &= E[(C * \varepsilon_k + v_k + u_k) * (C * \varepsilon_{k+1} + v_{k+1} + u_{k+1})^T] = \\
&= E[(C * \varepsilon_k + v_k + u_k) * (C * A * \varepsilon_k + C * G * w_k + v_{k+1} + u_{k+1})^T] = \\
&= E[C * A * \varepsilon_k * \varepsilon_k^T * C^T + v_k * v_{k+1}^T] = \\
&= C * A * P^- * C^T + E[v_k * (\lambda^T * v_k^T + \xi_k^T)] = \\
&= C * A * P^- * C^T + \lambda^T * R_v.
\end{aligned} \tag{7.12}$$

Kovariance výstupu pro okamžiky posunuté o dva časové indexy vypadá následovně:

$$\begin{aligned}
E[z_k^* * z_{k+2}^{*T}] &= E[(C * \varepsilon_k + v_k + u_k) * (C * \varepsilon_{k+2} + v_{k+2} + u_{k+2})^T] = \\
&= E \left[\begin{array}{c} (C * \varepsilon_k + v_k + u_k) * \\ * (C * A * A * \varepsilon_k + C * A * G * w_k + C * G * w_{k+1} + v_{k+2} + u_{k+2})^T \end{array} \right] = \\
&= E[C * A * A * \varepsilon_k * \varepsilon_k^T * C^T + v_k * v_{k+2}^T] = \\
&= C * A * A * P^- * C^T + E[v_k * (\lambda^T * (\lambda^T * v_k^T + \xi_k^T)^T + \xi_{k+1}^T)] = \\
&= C * A * A * P^- * C^T + \lambda^T * \lambda * R_v.
\end{aligned} \tag{7.13}$$

A takto bychom mohli pokračovat při výpočtu kovariance dále pomocí vztahu:

$$E[\widehat{z}_k^* * \widehat{z}_{k+j}^{*T}] = C * A^j * P^- * C^T + \lambda^j * R_v + \delta_{0j} * R_u, \tag{7.14}$$

kde index $j = 0, \dots, O$ značí časový posun jednoho výstupu od druhého a δ_{0j} je Kroneckerovo delta, jež je rovná jedné, pokud jsou si její dolní indexy rovny, a nule, pokud se indexy liší. Opět, jako v předchozím případě, figurují v (7.14) neznámé matice R a matice Q se ukrývá ve výpočtu ustálené hodnoty kovariance v rovnici (5.19).

Tohoto problému se ale zbavíme, pokud do rovnice (7.15) dosadíme ustálenou hodnotu kovariance z rovnice (7.7) a upravíme ji, jak je vidět v následujícím výpočtu:

$$E[z_k^* * z_{k+j}^{*T}]_S = (C \otimes C * A^j) * (I_{n^2 \times n^2} - A \otimes A)^{-1} * Q_S + \lambda^j * R_{v_S} + \delta_{0j} * R_{u_S}. \tag{7.15}$$

Pokud bychom rovnice (7.8), (7.9), (7.10), (7.11) převedli do tvaru (7.13) jako rovnici (7.12), lze přepsat rovnice do maticového tvaru, který utvoří následující nelineární rovnice:

$$\mathcal{A}(\theta(Q, R_u, R_v, \lambda)) = b. \tag{7.16}$$

Tato rovnice je oproti metodě ACLS nelineární. Z této rovnice lze navíc odhadnout parametry Q , R_u , R_v a λ , ale ne matici R_ξ . Tu lze vypočítat pomocí odhadů parametrů R_v a λ a rovnice (7.2) za pomoci následujícího vztahu:

$$cov[v_{k+1}] = cov[\lambda * v_k + \xi_k]. \tag{7.17}$$

Tato Ljapunovská rovnice má ustálené řešení v^- za předpokladu stabilního parametru λ . Proto lze tuto rovnici upravit na:

$$\begin{aligned} \text{cov}[v^-] &= \lambda * \text{cov}[v^-] * \lambda^T + \text{cov}[\xi_k], \\ R_v - \lambda * R_v * \lambda^T &= R_\xi, \end{aligned} \tag{7.18}$$

$$(I_{p^2 \times p^2} - \lambda \otimes \lambda) * R_{v_S} = R_{\xi_S}.$$

Matice \mathcal{A} v rovnici (7.14) vypadá tedy následovně:

$$\mathcal{A} = \begin{bmatrix} (C \otimes C) * (I_{n^2 \times n^2} - A \otimes A)^{-1} * Q_S + R_{v_S} + R_{u_S} \\ (C \otimes C * A) * (I_{n^2 \times n^2} - A \otimes A)^{-1} * Q_S + \lambda * R_{v_S} \\ \vdots \\ (C \otimes C * A^{O-1}) * (I_{n^2 \times n^2} - A \otimes A)^{-1} * Q_S + \lambda^{O-1} * R_{v_S} \end{bmatrix}. \tag{7.19}$$

V rovnici (7.17) je vektor pravé strany b tvořen kovariancemi inovační posloupnosti a má stejnou podobu jako (5.30). Lze ho také odhadnout pomocí vzorce (5.32) pouze s tou změnou, že za výstup z zvolíme alternativní výstup z^* . Odhad levé strany rovnice (7.17) bude tedy vypadat následovně:

$$\hat{b} = \begin{bmatrix} E[z_k^* * z_k^{*T}] \\ E[z_k^* * z_{k+1}^{*T}] \\ \vdots \\ E[z_k^* * z_{k+O-1}^{*T}] \end{bmatrix}_S. \tag{7.20}$$

Nyní lze vypočítat rovnici (7.17), kterou vyřešíme pomocí metody nelineárních nejmenších čtverců, a získáme nestranné odhady kovariančních matic působících jak na měření, tak i na rovnici měření. Je též vidět, že pokud bychom v této metodě zvolili λ a q rovné nule, dostali bychom se zpět k metodě ACLS, která by měla nulový zisk.

7.2 Vlastnosti metody

Jednou z velkých výhod této rozšířené metody ACLS je ta, že metoda je stále analyticky odvozená a dává nestranné odhady všech neznámých parametrů.

Nyní se zaměříme na nežádoucí vlastnosti, mezi které patří například volba zisku L . Pro zjednodušení byl zisk volen nulový, což značně zjednodušilo odvození. Metoda vedla na relativně jednoduché vztahy v nelineární rovnici (7.16). Tyto vztahy

nevedly k nutnosti použít složitější identifikační metodu pro nelineární rovnice a vystačili jsme si s nelineární metodou nejmenších čtverců.

Tato volba též ovlivnila množinu odhadovaných neznámých matic Q , R_u , R_ξ a λ , jež popisují vlastnosti poruch působící na systém. Při takto zvolených odhadovaných parametrech nelineární metoda nejmenších čtverců není robustní vůči počátečním podmínkám této metody a odhad často nekonvergoval ke správným maticím kvůli přílišné nelinearitě rovnice (7.16). Proto byly odhadovány matice Q , R_u , R_v a λ , pro které již nelineární metoda nejmenších čtverců byla robustní vůči počátečním podmínkám a dávala odhady, jež konvergovaly ke správným maticím.

7.3 Rozšířená metoda ACLS pro statický systém

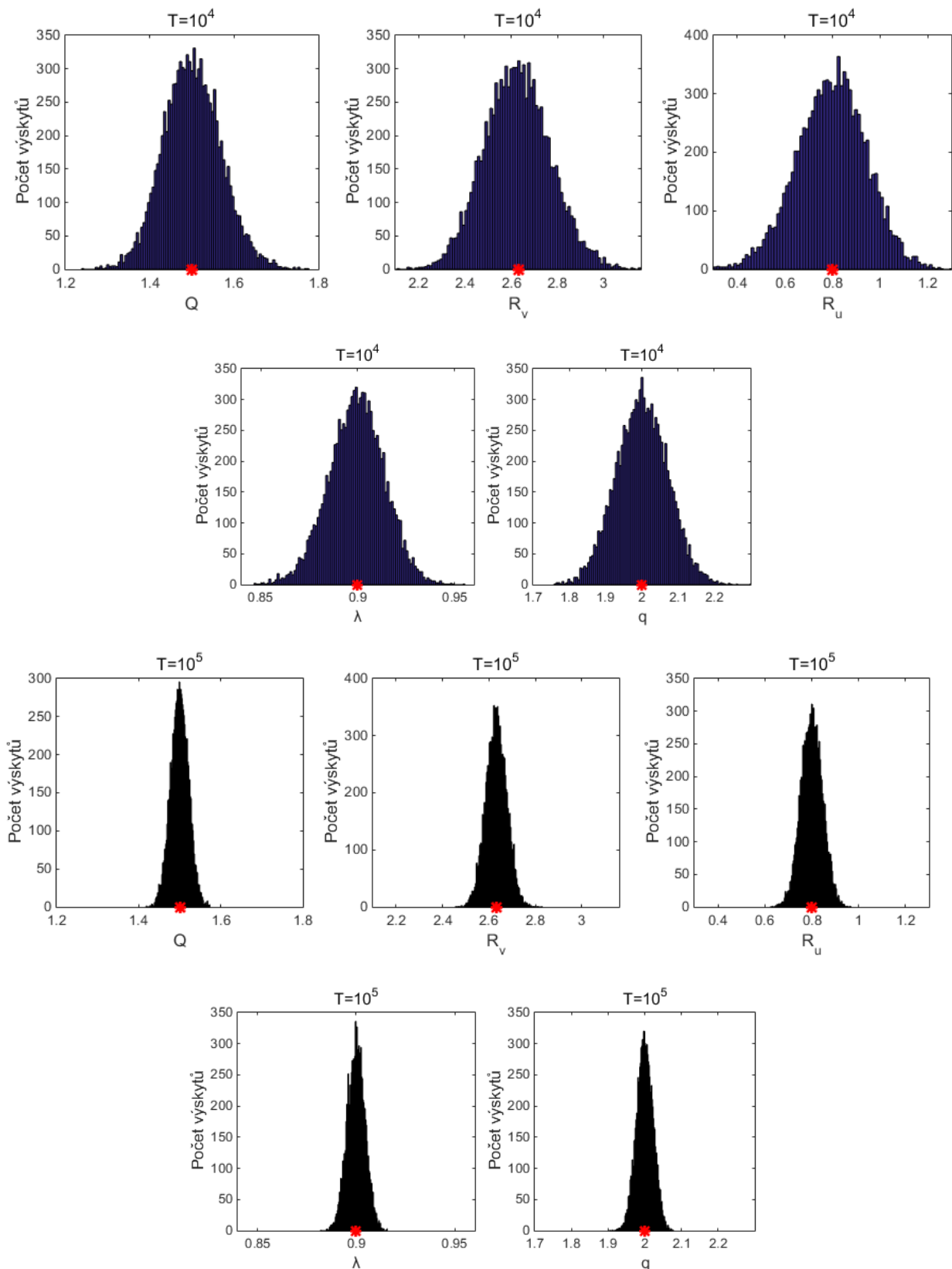
Pro odhad poruch působících na rovnici měření se využívá systémů bez dynamiky neboli statický systém. Takový systém uskutečníme například tím, že bychom měli jako měřící zařízení inerciální senzor, jenž by byl připevněn na desce, která by se nakláněla a otáčela do přesně změřených poloh, ve kterých by se provádělo měření. Díky tomu rovnice systému degeneruje tak, že matice dynamiky se změní na jednotkovou matici, a porucha působící na stav bude nulová. V takovémto případě budeme tedy moci získávat data ze senzoru, která budou ovlivněna pouze poruchami působícími na toto měření. Tento způsob se také využívá pro odhad poruch pomocí metody Allanovy variance jako v [23] – [26].

7.4 Numerická ilustrace

Na ukázkou metody odvozené v předchozí kapitole budou reprezentovány její výsledky. Pro první příklad bude zvolen tento dynamický systém:

$$A = -0.8; C = 1; Q = 1.5; R_\xi = 0.5; R_u = 0.8; \lambda = 0.9; q = 2; R_v \doteq 2.6316. \quad (7.21)$$

Tento dynamický systém byl simulován pro dvě různé délky měření, 10 000 Monte Carlo simulací a $O = 4$ kovarianční rovnice. Výsledné odhady byly vykresleny do následujících histogramů:



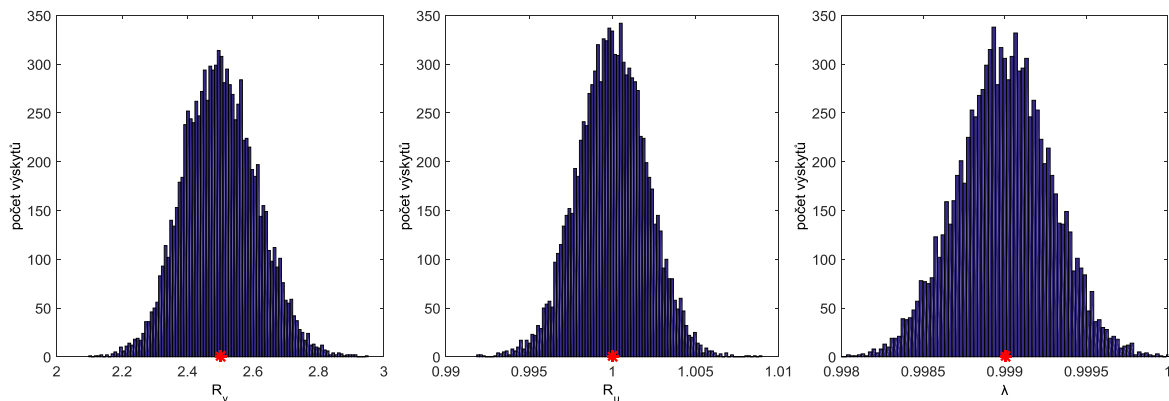
Obrázek 12 – Odhady rozšířené metody ACLS pro různá N , dynamický systém

kde na obrázku 12 jsou červeným křížkem vyznačeny skutečné hodnoty odhadovaného parametru. Z výsledku je tedy vidět, že tato metoda odhaduje parametry nestranné a variance tohoto odhadu se při zvětšení množiny dat měření snižuje.

Jako druhý příklad bude zvolen systém bez dynamiky, který vypadá následovně:

$$C = 1; R_{\xi} = 0.005; R_u = 1; \lambda = 0.999; q = 0; R_v \doteq 2.5013. \quad (7.22)$$

Tento systém byl nasimulován pro $N = 1\,000\,000$ vzorků měření, 10 000 Monte Carlo simulací a $O = 4$ kovarianční rovnice. Výsledky této simulace jsou vidět na následujících histogramech:



Obrázek 13 - Odhady modifikované metody ACLS, systém bez dynamiky

Je vidět, že i tyto odhady jsou nestranné stejně jako v předchozím případě.

Tento druhý příklad nebo spíše jeho výsledky se nám hodí jako protiklad k výsledkům, které jsme dostali od společnosti Honeywell, jež používá metodu Allanovy variance. Tato metoda se využívá pro kalibraci senzoru a především pro inerciální senzory, jak je diskutováno v [23] – [26]. Je navržena pro systém bez dynamiky s rovnicí měření se strukturou, jako je v rovnici (7.1), a parametry jako v rovnici (7.22). Využívá toho, že porucha působící na měření, se skládá ze dvou nezávislých poruch, z nichž jedna je korelovaná v čase a druhá porucha je bílá.

Metoda se dá použít i pro diskrétní skalární systém, jak bylo psáno v [27] a [28] následujícími kroky:

- i. Měření z_k o délce vzorků N je rozděleno do K úseků o velikosti M . Dále pak platí vzorce $K = \frac{N}{M}$ a $\tau_M = \frac{M}{f_S}$, kde τ_M je doba úseku a f_S je perioda vzorkování.
- ii. Pro každý úsek lze vypočítat odhad střední hodnoty pomocí vzorce

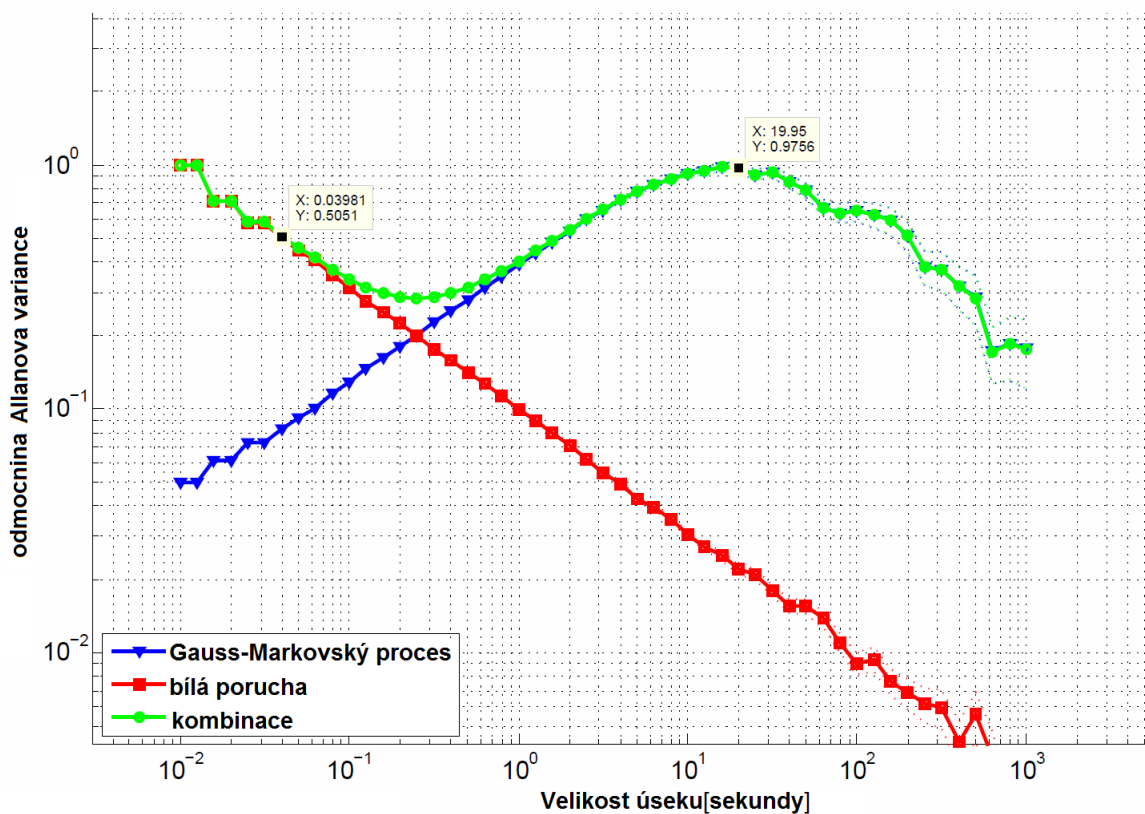
$$\hat{z}(M) = \frac{1}{M} \sum_{i=1}^M z_{(k-1)*M+i} \quad (7.23)$$

- iii. Hlavní částí Allanovy variance je výpočet samotné variance, kterou lze provést podle následujícího vzorce:

$$\sigma_A^2(\tau_M) = \frac{1}{2 * (K - 1)} \sum_{k=1}^{K-1} [\hat{z}_{k+1}(M) - \hat{z}_k(M)]^2 \quad (7.24)$$

- iv. Předchozí kroky se opakují pro různé délky M .

Výsledná Allanova variance se zobrazuje do logaritmického grafu, kde na ose x se vykresluje délka úseku v sekundách τ_M a na ose y se vykresluje odmocnina Allanovy variance $\sigma_A(\tau_M)$. Graf tedy pro systém (7.22), pokud vzorkovací frekvence f_S bude rovna 100 Hz, vypadá následovně:



Obrázek 14 – Odmocnina Allanovy variance v závislosti na délce úseku, logaritmické osy

Na obrázku 14 je vidět působení bílého, Gauss-Markovského procesu a jejich kombinace.

Samotný bílý šum se projevuje jako řada se sklonem -0.5 na logaritmické stupnici na obou osách. Každý bod této řady reprezentuje Allanovu varianci bílého šumu, ze které můžeme vypočítat varianci tohoto bílého šumu pomocí vzorce:

$$R_u = \sigma_A^2(\tau_M) * \tau_M * f_s, \quad (7.25)$$

kde bod τ_M reprezentuje velikost úseku v sekundách z obrázku 14, ve kterém se začíná kombinace obou variancí a bílé poruchy vzdalovat. Tento bod je v grafu vyznačený.

Časově korelovaný šum neboli Gauss-Markovského proces zobrazený v obrázku 14 vytváří „kopeček“, u kterého měříme vrchol. Díky tomuto vrcholu lze vypočítat ustálenou hodnotu variance Gauss-Markovského procesu R_v a konstantu λ pomocí následujících vzorců:

$$\tau_c = \frac{\tau_{max}}{1.89} \quad (7.26)$$

$$R_v = \left(\frac{\sigma_A(\tau_{max})}{0.62} \right)^2, \quad (7.27)$$

kde bod τ_c reprezentuje korelaci v čase, τ_{max} reprezentuje velikost úseku, při které dochází ke zlomu v obrázku 14 a vytvoření „kopečku“, a $\sigma_A^2(\tau_{max})$ reprezentuje varianci v bodě τ_{max} . Tento bod je v grafu také vyznačený.

Pokud bychom tedy odměřili hodnoty z obrázku 14 a hodnoty, které máme zadané, zjistili bychom, že zatím máme tyto hodnoty:

$$f_s = 100 [Hz]; \tau_M = 0.03981[s]; \sigma_A(\tau_M) = 0.5051[s] \quad (7.28)$$

$$\tau_{max} = 19.95[s]; \sigma_A(\tau_{max}) = 0.9756.$$

Z těchto hodnot bychom mohli pomocí vzorce (7.25) vypočítat, že $R_u = 1.0157$. Dále můžeme zjistit pomocí vzorce (7.27), že $R_v = 2.4761$, a pomocí vzorce (7.26), že časová konstanta je $\tau_c = 10,5$. Časová konstanta se dá převést do diskretní oblasti a vypočítat pomocí ní konstantu λ následujícím vzorcem:

$$\lambda = e^{\frac{-1}{\tau_c * f_s}}. \quad (7.29)$$

Pomocí tohoto vzorce bylo tedy vypočteno $\lambda = 0.9991$ a pomocí vzorce (7.18) $R_\xi = 0.0047$.

Pokud bychom tedy výsledky z druhého příkladu počítaného jak pomocí modifikované metody ACLS, tak i metody Allanovy variance dali do tabulky, výsledky by vypadaly následovně:

	Skutečné parametry	Allanova variance	Rozšířená ACLS
R_u	1	1.0157	1.0000
λ	0.999	0.9991	0.9990
R_v	2.5013	2.4761	2.5013
R_ξ	0.005	0.0047	0.0050

Tabulka 1 – Výsledky metody Allanovy variance a rozšířené metody ACLS

Je vidět, že modifikovaná metoda ACLS dává nestranné odhady parametrů, jejichž variance se s přibývajícím počtem dat snižuje. Zato metoda Allanovy variance nemá oproti modifikované ACLS tak dobré výsledky, ale i přesto jsou odhady velice blízko skutečným parametrům.

8. Softwarová implementace metody ACLS

Tato kapitola se zabývá programem využívající metodu ACLS pro odhad kovariančních matic poruch a jednotlivými volbami tohoto programu.

8.1 Obecné informace

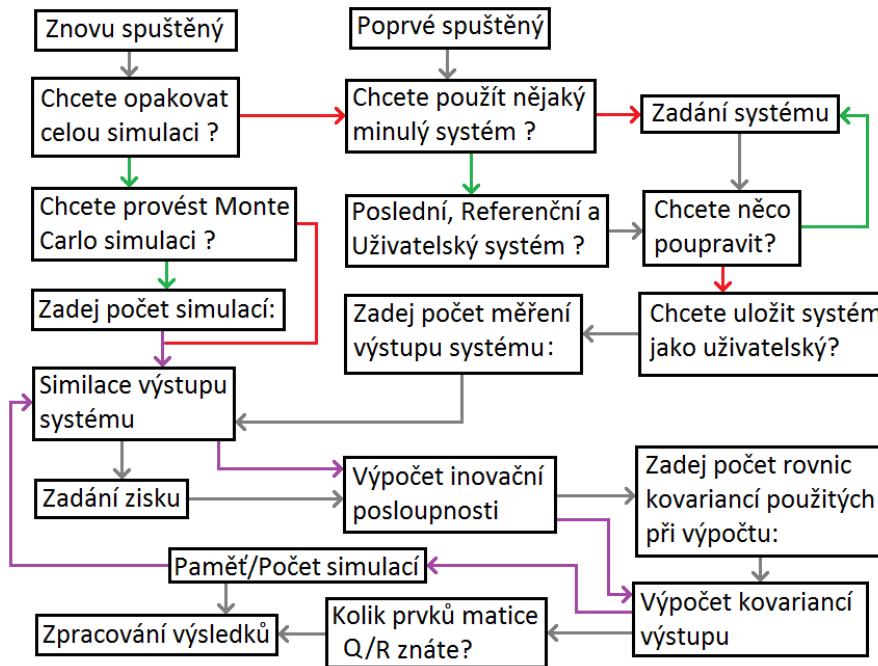
Metoda ACLS byla implementována v programu Matlab. Výsledný program se skládá ze souboru *aclsmethod.m*, jenž je hlavním jádrem programu, a dále souborů *ACLS_11.mat*, *ACLS_12.mat*, *ACLS_13.mat*, *ACLS_21.mat*, ... , *ACLS_33.mat*, které ukrývají referenční modely a slovníky *cz.mat* a *en.mat*.

Metoda je navržena pomocí dvou reprezentací metody ACLS v [3] a [21]. Volba výpočtu mezi těmito dvěma reprezentacemi se vždy provádí na počátku simulace, kdy, pokud chceme první metodu z [3], stačí do proměnné *method* dát číslo jedna, a pokud chceme druhou metodu z [21], do proměnné dáme číslo dva.

Další volbou, jež se provádí na začátku, je volba jazyka, ve kterém program probíhá. Na výběr jsou dva jazyky – český a anglický. Volba se provádí načtením slovníku metodou *load('soubor.mat')*, pokud místo *souboru.mat* je zadáno *cz.mat*, program bude v českém jazyce, a pokud je zadáno *en.mat*, program bude v anglickém jazyce.

8.2 Postup a možnosti zadávání

Na úvod bude ukázán obrázek, jenž ukazuje jednotlivé hlavní části, volby a průchody programem, který vypadá následovně:



Obrázek 15 - Postup programem

Jednotlivé bloky ukazují, co se v dané části děje nebo na co se bude uživatele program ptát, šedé čáry ukazují přirozený postup programem, zelená čára značí odpověď *ano* na otázku, červená čára značí odpověď *ne* na otázku a fialová čára značí postup metody Monte Carlo.

Nyní si blíže popíšeme některé bloky a postupy funkcí. Jako první se zaměříme na volbu *Poslední, Referenční a Uživatelský systém?*, kde na počátku systému nemáme ještě žádný poslední ani uživatelský systém, a tak pokud v bloku *Chcete použít nějaký minulý systém?* zvolíme volbu *ano*, bude rovnou nabídnuto z devíti různých referenčních systémů. Dalším popisovaným blokem je blok *Zadání systému*. Zde se dá zadat kompletně celý systém pro metodu ACLS, ať už volba matice systému, která se dá zvolit přímo či přes zadání vlastních čísel matice, nebo například počátek stavu systému. V bloku *Zadání zisku* se dá zvolit zisk L buď po prvcích, nebo pomocí matic Q , R a vzorců (3.14) a (3.15). Pokud v bloku *Chcete provést Monte Carlo simulaci?* zvolíme *ano*, budeme dále zadávat počet Monte Carlo simulací a výsledkem nebude pouze odhad matic Q a R , ale bude také k dispozici i variance těchto odhadů. Pokud zvolíme *ne*, bude odhad neznámých matic pouze opakován s nastavením, jež bylo zvoleno naposledy. Posledním popisovaným je blok

Kolik prvků matice Q/R znáte?, ve kterém se dají vybrat prvky, jež jsou známé, a zbylé neznámé unikátní prvky matic, které budou odhadnuty.

Při jakémkoliv zadávání do programu je kontrolováno, zda je zadán správný formát čísel či textu. Při nesprávném zadání program upozorní uživatele a vyžádá si nové zadání. Také se kontroluje například při zadání zisku, zda je rovnice (6.9) stabilní nebo zda je zvoleno dostatečné množství rovnic pro řešení metody nejmenších čtverců. V jakémkoliv zadávání lze též zadat písmeno *e*, které ukončí program.

Výsledný odhad pomocí programu vypadá například následovně:

```
Hodnost matice Anc - 6 | Pocet odhadovanych prvku - 6
Delka matice Anc - 6
```

```
Odhad varianci Q
skutecna | odhadnuta
  1      0.08 | 0.98275   0.068123
0.08     2 | 0.068123   2.0001
```

```
Vlastni cisla matice Q
0.99364 | 0.97821
 2.0064 | 2.0046
```

```
Variance odhadu Q | soucet vseh prvku
0.044494   0.12444 |
 0.12444   0.19159 | 0.48496
```

```
Kvadraticky rozdil vseh prvku od skutecne matice Q
0.00057961
```

```
Odhad varianci R
skutecna | odhadnuta
  5      0.3 | 4.9936   0.29408
0.3     2 | 0.29408   2.0002
```

```
Vlastni cisla matice R
1.9703 | 1.9716
5.0297 | 5.0223
```

```
Variance odhadu R | soucet vseh prvku
 0.26247   0.033127 |
0.033127   0.024735 | 0.35346
```

```
Kvadraticky rozdil vseh prvku od skutecne matice R
0.00011051
```

Obrázek 16 - Ukázka výsledku odhadu z programu

9. Zhodnocení

Tato diplomová práce se zabývala odhadem vlastností poruch působících na systém popsaný stavovou reprezentací.

V úvodu této diplomové práce bylo ukázáno, jak lze model převést do stavového popisu a jak na tento popis působí poruchy. Dále bylo ukázáno, jak pomocí Kalmanova filtru odhadnout stav systému a jaké jsou k tomuto odhadu předpoklady, zejména pak znalost popisu poruch. Proto byla představena autokovarianční metoda pro odhad kovariančních matic poruch, jež využívá k odhadu inovační posloupnosti lineárního filtru. Tato metoda má ty pozitivní vlastnosti, že je analyticky odvozená a že poskytuje nestranné odhady kovariančních matic poruch. Metoda má také i slabiny. Jednou z těchto slabin je například nutnost mít dostatečné množství senzorů, aby metoda dokázala odhadnout všechny prvky kovariančních matic. Proto byl v této práci ukázán výpočet maximálního počtu odhadnutelných prvků. Další slabinou je to, že metoda je založena na uživatelské volbě zesílení lineárního filtru a výsledná kvalita tohoto odhadu je závislá na tomto zesílení. Proto byl ukázán výpočet, díky kterému uživatel nemusí volit zisk libovolně, ale tak, aby výsledný odhad byl kvalitní. Také byla ukázána možnost rozšíření metody pro časově korelované poruchy působící na měření, jako například u inerciálních senzorů, a byla ukázána i numerická ilustrace dosažených výsledků. Také byla ukázána kvalita odhadu této rozšířené metody oproti metodě založené na Allanově varianci, jež se standardně využívá pro odhad poruch působících na inerciální senzory. Na závěr byl představen program vytvořený v prostředí Matlab, který byl použit při testování a numerických ilustracích této autokovarianční metody.

Hlavní cíl této diplomové práce i všechny její dílčí cíle byly splněny.

Příloha

Tato kapitola se zabývá metodami lineárních a nelineárních nejmenších čtverců a Kroneckerovo algebrou.

I. Metoda lineárních nejmenších čtverců

Tato estimační metoda se používá k nalezení neznámých parametrů lineární rovnice. Těchto rovnic může být několik různých, v tomto případě bude metoda řešit problém přeurených rovnic, při které hledáme parametry jedné rovnice, které s nejmenší kvadratickou chybou odpovídají všem rovnicím. Další možností, kterou se budeme zabývat, je ta, že máme množinu měření, kterou generuje jedna lineární rovnice, na níž působí bílá porucha se střední hodnotou nula, jako bylo diskutováno v [20]. Pro příklad zavedeme lineární rovnici, která generuje měření a vypadá následovně:

$$\mathcal{A} * \theta + e = \hat{b}, \quad (9.1)$$

kde vektor θ značí vektor neznámých parametrů, které se snažíme odhadnout, vektor e značí neznámou bílou poruchu se střední hodnotou nula, která vzniká buď působením neurčitosti měření, nebo nepřesností parametru θ , a vektor \mathcal{A} je známý vektor regresorů. Rovnice se dá přepsat do následujícího tvaru:

$$e = \mathcal{A} * \theta - \hat{b}. \quad (9.2)$$

Z této rovnice je dobře viditelné, že pokud bychom chtěli zjistit parametry rovnice, musíme minimalizovat poruchu na levé straně tím, že se budeme snažit najít takové parametry θ , pro které bude chyba e minimální. Hledáme tak řešení následující funkce:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}}(e^T * e). \quad (9.3)$$

Pro nalezení řešení této funkce je dobré nejdříve rozepsat součin $e^T * e$, který vypadá následovně:

$$\begin{aligned} e^T * e &= (\mathcal{A} * \theta - \hat{b})^T * (\mathcal{A} * \theta - \hat{b}) = \\ &= \theta^T * \mathcal{A}^T * \mathcal{A} * \theta + \hat{b}^T * \hat{b} - \theta^T * \mathcal{A}^T * \hat{b} - \hat{b}^T * \mathcal{A} * \theta. \end{aligned} \quad (9.4)$$

Nyní pro nalezení argumentu minima funkce (9.3) musíme derivovat rovnici (9.4) podle parametru θ , jak je vidět v následujícím výpočtu:

$$\frac{\partial(e^T * e)}{\partial \theta} = 2 * \mathcal{A}^T * \mathcal{A} * \hat{\theta} - 2 * \mathcal{A}^T * \hat{b}. \quad (9.5)$$

Tato derivace se musí rovnat nule, pokud chceme nalézt minimum funkce (9.3), jak je vidět v následujícím výpočtu:

$$\begin{aligned} 2 * \mathcal{A}^T * \mathcal{A} * \hat{\theta} - 2 * \mathcal{A}^T * \hat{b} &= 0 \\ \mathcal{A}^T * \mathcal{A} * \hat{\theta} - \mathcal{A}^T * \hat{b} &= 0 \\ \hat{\theta} &= (\mathcal{A}^T * \mathcal{A})^{-1} * \mathcal{A}^T * \hat{b} = \mathcal{A}^\dagger * \hat{b}, \end{aligned} \quad (9.6)$$

kde matici \mathcal{A}^\dagger se také říká pseudoinverze.

Provedeme výpočet nestrannosti tohoto odhadu, který vypadá následovně:

$$\begin{aligned} E[\hat{\theta}] &= E[(\mathcal{A}^T * \mathcal{A})^{-1} * \mathcal{A}^T * \hat{b}] = \\ &= E[(\mathcal{A}^T * \mathcal{A})^{-1} * \mathcal{A}^T * (\mathcal{A} * \theta + e)] = \\ &= E[(\mathcal{A}^T * \mathcal{A})^{-1} * \mathcal{A}^T * \mathcal{A} * \theta] + E[(\mathcal{A}^T * \mathcal{A})^{-1} * \mathcal{A}^T * e] = \\ &= E[\theta] + (\mathcal{A}^T * \mathcal{A})^{-1} * \mathcal{A}^T * E[e] = \theta. \end{aligned} \quad (9.7)$$

Z tohoto výpočtu je vidět, že se střední hodnota odhadu rovná skutečným parametrům, a tudíž je odhad nestranný. Toto platí pouze pokud e je bílá porucha se střední hodnotou nula.

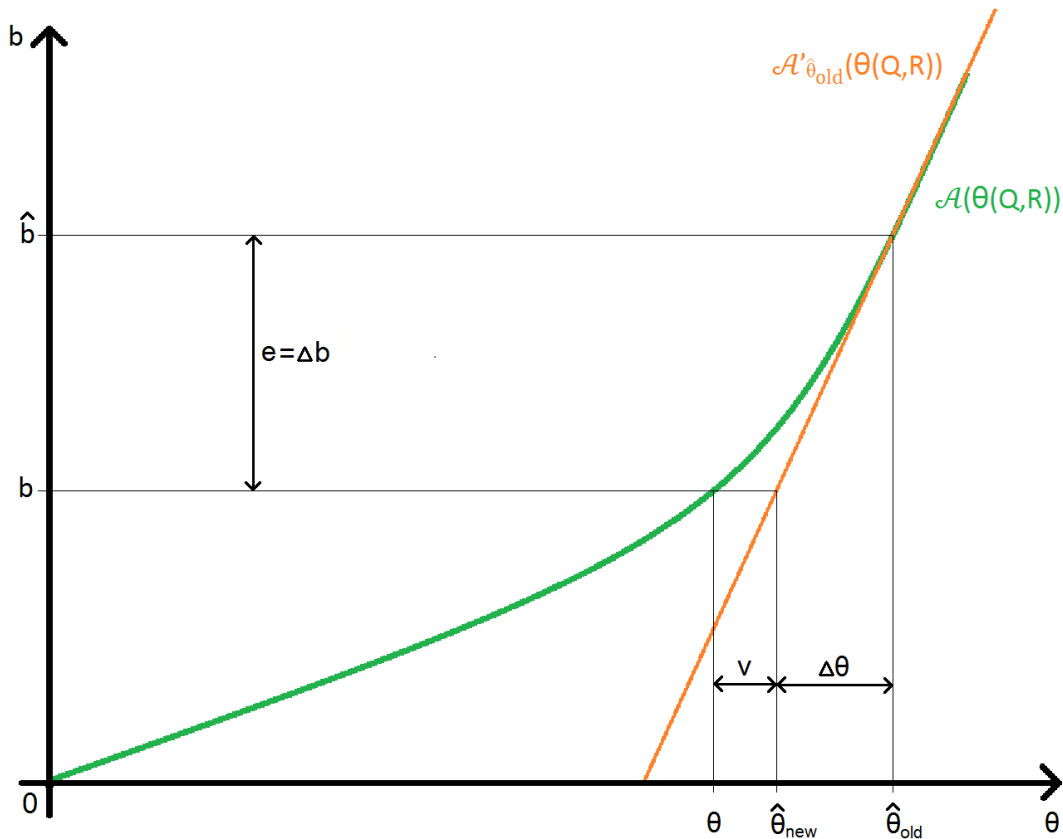
II. Metoda nelineárních nejmenších čtverců

Tato estimační metoda se používá oproti předchozí metodě již podle názvu k nalezení neznámých parametrů pro nelineární rovnice. Stejně jako v předchozí metodě se budeme zabývat pouze případy, kdy máme množinu měření, kterou generuje jedna nelineární rovnice, na níž působí bílá porucha se střední hodnotou nula.

Pro příklad zavedeme nelineární rovnici, která generuje měření a vypadá následovně:

$$\mathcal{A}(\theta) + e = \hat{b}. \quad (9.8)$$

Pro jednodušší vysvětlení metody bude nejprve ukázán obrázek, na kterém je tato metoda prezentovaná.



Obrázek 17 – Znázorněný jeden krok metody nelineárních nejmenších čtverců

Metoda začíná tak, že se nejprve zvolí počáteční parametr rovnice $\hat{\theta}_{old}$.

Poté metoda pokračuje následujícími kroky:

- i. Nelineární funkce $\mathcal{A}(\theta)$ se linearizuje v bodě $\hat{\theta}_{old}$ výpočtem

$$\mathcal{A}'_{\hat{\theta}_{old}} = \frac{\partial \mathcal{A}(\theta)}{\partial \hat{\theta}_{old}}, \quad (9.9)$$

kde $\mathcal{A}'_{\hat{\theta}_{old}}$ značí linearizaci funkce $\mathcal{A}(\theta)$ v bodě $\hat{\theta}_{old}$.

- ii. V dalším kroku provedeme měření b , ze kterého můžeme díky linearizaci v předchozím kroku a obrázku 16 odvodit následující vztahy:

$$\mathcal{A}'_{\hat{\theta}_{old}} * \hat{\theta}_{old} = \mathcal{A}(\theta) + e = \hat{b} \quad (9.10)$$

$$\mathcal{A}'_{\hat{\theta}_{old}} * \hat{\theta}_{new} = \mathcal{A}(\theta) = b. \quad (9.11)$$

- iii. Díky vztahům (9.10), (9.11) a znalosti obrázku 12 lze vypočítat chybu e neboli Δb pomocí následujícího vztahu:

$$\begin{aligned} e = \hat{b} - \mathcal{A}(\theta) &= \mathcal{A}'_{\hat{\theta}_{old}} * \hat{\theta}_{old} - \mathcal{A}'_{\hat{\theta}_{old}} * \hat{\theta}_{new} = \\ &= \hat{b} - b = \Delta b. \end{aligned} \quad (9.12)$$

V tomto kroku při výpočtu $\mathcal{A}'_{\hat{\theta}_{old}} * \hat{\theta}_{new}$ se dopouštíme chyby $v = \hat{\theta}_{new} - \theta$, která vzniká důsledkem linearizace funkce $\mathcal{A}(\theta)$ v bodě $\hat{\theta}_{old}$, kterou jsme ovšem použili v bodě $\hat{\theta}_{new}$.

- iv. Dále se vypočítá za pomoci vztahu (9.12) změna $\Delta\theta$, díky které se přiblížíme ke skutečnému parametru θ pomocí následujícího vztahu:

$$\mathcal{A}'_{\hat{\theta}_{old}} * (\hat{\theta}_{old} - \hat{\theta}_{new}) = \Delta b \quad (9.13)$$

$$\Delta\theta = \hat{\theta}_{old} - \hat{\theta}_{new} = \mathcal{A}'_{\hat{\theta}_{old}}^{-1} \Delta b.$$

- v. Díky znalosti $\Delta\theta$ vypočítané v předchozím kroku lze vypočítat $\hat{\theta}_{new}$, který je oproti $\hat{\theta}_{old}$ blíže skutečnému parametru θ pomocí následujícího vzorce:

$$\hat{\theta}_{new} = \hat{\theta}_{old} + \Delta\theta \quad (9.14)$$

- vi. Posledním krokem je následné nahrazení:

$$\hat{\theta}_{old} = \hat{\theta}_{new} \quad (9.15)$$

a následný návrat do kroku i, dokud nebude rozdíl $\Delta\theta$ malý.

Tyto kroky za předpokladu, že použijeme vhodnou funkci a počáteční podmínku, budou generovat odhad $\hat{\theta}_{new}$, který se bude postupně přibližovat skutečné hodnotě θ , a chyba v se bude blížit nule.

III. Kroneckerova algebra

Kroneckerova algebra zavádí do lineární algebry vzorce a pravidla, které umožňují provádět složité výpočty jednodušeji. Proto zde bude ukázáno několik těchto pravidel, která byla diskutována v [18] s [29].

Pro příklad zavedeme následující matice:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}; B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix}; C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}, \quad (9.16)$$

kde matice A má dimenzi $n \times m$, matice B má dimenzi $p \times q$ a matice C má dimenzi $r \times t$.

Jako první si ukážeme vektorový operátor vec , který aplikovaný na matici A udělá toto:

$$vec(A) = [a_{11} \ a_{21} \ a_{31} \ a_{12} \ a_{22} \ a_{32} \ a_{13} \ a_{23} \ a_{33}]^T. \quad (9.17)$$

Dalším je takzvaný Kroneckerův součin, který aplikovaný na matice A a B vypadá následovně:

$$A \otimes B = \begin{bmatrix} a_{11} * B & a_{12} * B & a_{13} * B \\ a_{21} * B & a_{22} * B & a_{23} * B \\ a_{31} * B & a_{32} * B & a_{33} * B \end{bmatrix}. \quad (9.18)$$

U tohoto součinu je výsledná dimenze $n * p \times m * q$.

Následující vztah se nazývá Kroneckerův součet, počítá se pouze pro čtvercové matice, a to takto:

$$A \oplus B = A \otimes I_{p \times q} + I_{n \times m} \otimes B. \quad (9.19)$$

U tohoto součinu je výsledná dimenze $n * p \times m * q$.

Posledními zde zmíněnými z mnoha dalších jsou tyto vztahy:

$$vec(A * C * B) = (B^T \otimes A) * vec(C) \quad (9.20)$$

$$vec(B * A) = (A^T \otimes I_{m \times p}) * vec(B), \quad (9.21)$$

kde pokud platit $n = q$, $m = r$ a $t = p$, má rovnice (9.20) výslednou dimenzi rovnou $n^2 \times 1$ a rovnice (9.21) má výslednou dimenzi rovnou $m^2 \times 1$.

Reference

- [1] R. K. Mehra, “On the identification of variances and adaptive filtering”, *IEEE Transactions on Automatic Control*, vol. 15, no. 2, pp. 175–184, 1970.
- [2] P. R. Bélanger, “Estimation of noise covariance matrices for a linear time-varying stochastic process,” *Automatica*, vol. 10, no. 3, pp. 267–275, 1974.
- [3] B. J. Odelson, M. R. Rajamani, and J. B. Rawlings, “A new autocovariance least-squares method for estimating noise covariances,” *Automatica*, vol. 42, no. 2, pp. 303–308, 2006.
- [4] B. M. Åkesson, J. B. Jørgensen, N. K. Poulsen, and S. B. Jørgensen, “A generalized autocovariance least-squares method for Kalman filter tuning,” *Journal of Process Control*, vol. 18, no. 7–8, pp. 769–779, 2008.
- [5] M. Šimandl and J. Duník, “Estimation of noise covariance matrices for periodic systems,” *International Journal of Adaptive Control and Signal Processing*, vol. 25, pp. 928–942, 2011.
- [6] D. G. Lainiotis, “Optimal adaptive estimation: Structure and parameters adaptation,” *IEEE Transactions On Automatic Control*, vol. 16, no. 2, pp. 160–170, 1971.
- [7] R. L. Kashyap, “Maximum likelihood identification of stochastic linear systems,” *IEEE Transactions on Automatic Control*, vol. 15, no. 1, pp. 25–34, 1970.
- [8] K. A. Myers and B. D. Tapley, “Adaptive sequential estimation with unknown noise statistics,” *IEEE Transactions on Automatic Control*, vol. 21, no. 8, pp. 520–523, 1976.
- [9] S. Verdú and H. V. Poor, “Minimax linear observers and regulators for stochastic systems with uncertain second-order statistics,” *IEEE Transactions On Automatic Control*, vol. 29, no. 6, pp. 499–511, 1984.
- [10] P. van Overshee and B. de Moor, *Subspace Identification for Linear Systems: Theory-Implementation-Applications*. Kluwer Academic Publishers, 1996.
- [11] L. Ljung, *System Identification: Theory for the User*. UpperSaddle River, NJ: Prentice-Hall, 1999.
- [12] D. M. Wiberg, T. D. Powell, and D. Ljungquist, “An on-line parameter estimator for quick convergence and time-varying linear systems,” *IEEE Transactions on Automatic Control*, vol. 45, no. 10, pp. 1854–1863, 2000.

- [13] S. Särkkä and A. Nummenmaa, "Recursive noise adaptive Kalman filtering by variational bayesian approximations," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 596–600, 2009.
- [14] B. Carew and P. R. Bélanger, "Identification of optimum steadystate gain for systems with unknown noise covariances," *IEEE Transactions on Automatic Control*, vol. 18, no. 6, pp. 582–587 s., 1973.
- [15] R. Bos, X. Bombois, and P. M. J. Van den Hof, "Designing a Kalman filter when no noise covariance information is available," in *Proceedings of the 16th IFAC World Congress, Prague, Czech Republic, July 2005*.
- [16] R. K. Mehra, "Approaches to adaptive filtering," *IEEE Transactions on Automatic Control*, vol. 17, no. 10, pp. 693–698, 1972.
- [17] J. Duník, M. Šimandl, and O. Straka, "Methods for estimating state and measurement noise covariance matrices: Aspects and comparison," in *Proceedings of 15th IFAC Symposium on System Identification, Saint-Malo, France, July 2009*.
- [18] J. W. Brewer, "Kronecker products and matrix calculus in system theory", *IEEE Transactions*
- [19] J. Melichar, *Lineární systémy*. Plzeň: Západočeská univerzita, 117s. 2011.
- [20] M. Šimandl, *Identifikace systémů a filtrace*. Plzeň: Západočeská univerzita, 184s., 2011.
- [21] M. R. Rajamani, J. B. Rawlings, "Estimation of the disturbance strukture from data using semidifinite programming and optimal weighting" *Automatica*, vol. 45, pp. 142–148, 2009.
- [22] F. V. Lima, M. R. Rajamani, T. A. Soderstrom, a J. B. Rawlings, „Covariance and State Estimation of Weakly Observable Systems: Application to Polymerization Processes“, *IEEE Transactions on Automatic Control*, vol. 21, no. 3, pp. 1249-1257, 2013.
- [23] R. M. Rogers, *Applied Mathematics in Integrated Navigation Systems (2nd Edition)*. AIAA, 2003
- [24] J. A. Farrell, *Aided Navigation - GPS with High Rate Sensors*. McGraw Hill, 2008.
- [25] "IEEE standard specification format guide and test procedure for single-axis interferometric fiber optic gyros," *IEEE Std. 952-1997*, 2008.
- [26] "IEEE standard definitions of physical quantities for fundamental frequency and time metrology - random instabilities," *IEEE Std. 952-1997*, 1999.

[27] L. C. Ng and D. J. Pines, "Characterization of ring laser gyro performance using the allan variance method," *Journal of Guidance, Control, and Dynamics*, vol. 20, no. 1, pp. 211–214, 1997.

[28] M. Soták, "Determining stochastic parameters using an unified method", *Acta Electrotechnica et Informatica*, vol. 9, no. 2, pp. 59–63, 2009.

[29] P. A. Regalia, S. K. Mitra, "Kronecker Products, Citary Matrices and Signal Processing Applications" *Siam review*, vol. 31, no. 4, pp. 586–613, 1989.

[30] Vladimír Havlena, Jan Štecha, "Moderní teorie řízení", VÚT, pp. 234-236, 1999.