

Hodnocení kvality odhadu stavu stochastických systémů

Jindřich Havlík

16. května 2014

Západočeská univerzita v Plzni

Fakulta aplikovaných věd

Katedra kybernetiky

DIPLOMOVÁ PRÁCE

PLZEŇ, 2014

JINDŘICH HAVLÍK

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 16. května 2014

.....
vlastnoruční podpis

Poděkování

Chtěl bych poděkovat Ing. Ondřeji Strakovi, Ph.D., za odborné vedení práce a poskytnutí cenných rad.

Anotace

Diplomová práce se zabývá hodnocením kvality odhadu stavu nelineárních stochastických dynamických systémů diskrétních v čase. Je hodnocena kvalita bodových odhadů stavu poskytovaných lokálními estimátory, jejichž struktura odpovídá tzv. gaussovskému filtru. Práce představuje metriky absolutní chyby odhadu, relativní chyby odhadu, míry četnosti, ale i skalární hodnocení kovarianční matice chyby odhadu, metriky věrohodnosti a Cramér-Raovu mez. Vlastnosti metrik jsou ilustrovány na konkrétních problémech z praxe a shrnuty do přehledných tabulek. Na závěr je představena metodika dávající návod k řešení problému hodnocení kvality odhadu stavu.

Klíčová slova: Odhad stavu, hodnocení kvality, metriky, věrohodnost odhadu, gaussovský filtr

Abstract

The diploma thesis deals with performance evaluation of state estimates of nonlinear stochastic dynamic systems. The thesis focuses on the performance measures of point estimates generated by local estimators, which have the structure of Gaussian Filter. The thesis presents absolute error metrics, relative error metrics, frequency count measures, scalar matrix measures, credibility measures, and also Cramér-Rao bound. Properties of performance measures are summed up in the table and illustrated in a few practical problems. Finally, a recommendation on how to proceed when measuring performance is given.

Keywords: State estimation, performance measures, metrics, credibility, Gaussian Filter

Obsah

1	Úvod	5
2	Úloha odhadu stavu	7
2.1	Popis systému	7
2.2	Formulace a řešení úlohy odhadu	8
2.3	Globální metody odhadu stavu	9
2.4	Lokální metody odhadu stavu	9
2.4.1	Rozšířený Kalmanův filtr	12
2.4.2	Unscentovaný Kalmanův filtr	13
3	Cíle práce	15
4	Metriky chyby odhadu	17
4.1	Metriky absolutní chyby	17
4.1.1	Odmocnina ze střední kvadratické chyby - RMSE	18
4.1.2	Průměrná eukleidovská chyba - AEE	18
4.1.3	Harmonický průměr chyb - HAE	19
4.1.4	Geometrický průměr chyb - GAE	19
4.1.5	Medián a modus chyby	20
4.1.6	Iterační střední rozsah chyby - IMRE	20
4.1.7	Zobecnění metrik a poznámky	20
4.2	Metriky relativní chyby	21
4.2.1	Relativní verze vybraných metrik absolutní chyby	22
4.2.2	Kvocient chyby bayesovského odhadu - BEEQ	23
4.2.3	Chyba odhadu vztážená k chybě měření - EMER	24
4.2.4	Redukce chyby bayesovského odhadu - BERF	24
4.2.5	Poznámky	25
4.3	Míry četnosti	25
4.3.1	Míra úspěšnosti	26
4.3.2	Míra neúspěšnosti	26
4.3.3	Poznámky	26
4.4	Problém fyzikálně různorodých stavů a stavů řádově odlišných hodnot	27
5	Metriky kovarianční a MSE matice	28
5.1	Hodnocení samostatné matice	28
5.2	Věrohodnost odhadu	30
5.2.1	Jednoduchá hodnocení věrohodnosti	30

<i>OBSAH</i>	4
5.2.2 Normalizovaný kvadrát chyby odhadu - NEES	31
5.2.3 Index nevěrohodnosti a indikátor inklinace - NCI a I^2	32
5.3 Cramér-Raova mez - porovnání MSE s mezí poznatelnosti	33
6 Metodika hodnocení kvality odhadu	35
7 Aplikace metrik ve vybraných úlohách	39
7.1 Odhad polohy pohybujícího se objektu	39
7.2 Odhad veličiny chemického procesu	44
7.3 Nestacionární růstový model používaný v ekonometrii	47
8 Závěr	50

Kapitola 1

Úvod

Diplomová práce se zabývá hodnocením kvality odhadu stavu dynamických stochastických systémů. Cílem úlohy odhadu (estimace) stavu stochastických systémů je na základě měřených dat a jisté apriorní informace o sledovaném systému navrhnout co nejlepší odhad neměřitelné složky stavu systému. Jako podmnožinu obecné úlohy odhadu stavu lze chápat i estimaci parametrů nebo též parametrickou identifikaci. Podle vzájemného vztahu časového okamžiku odhadovaného stavu a časového okamžiku měření se dělí úloha odhadu stavu na predikci, filtraci a vyhlazování. Na základě měření až do současného časového okamžiku hledá úloha predikce odhad budoucích stavů, úloha filtrace hledá odhad stavu v současném okamžiku a úloha vyhlazování hledá odhad minulých stavů.

Moderní teorie odhadu má své kořeny ve 40. letech minulého století, kdy N. Wiener, označovaný za otce kybernetiky, navrhl frekvenční přístup k řešení lineárního filtračního problému spojitého v čase. Pro stejný problém, ale tentokrát v diskrétním čase, představil kolem roku 1960 R. E. Kalman stavový přístup, na němž je založen známý algoritmus Kalmanova filtru. Následný rapidní rozvoj nelineární filtrace od konce 60. let minulého století úzce souvisel s tehdejšími závody ve zbrojení (navigační systémy balistických hlavic), a také s vesmírnými závody (program Apollo). Dodnes je hlavní pole působnosti úlohy odhadu stavu v oblasti zpracování signálů, detekce poruch, navigace a sledování pohybujících se objektů. Úloha odhadu tak má skutečně široké uplatnění v mnoha oborech lidské činnosti: například v letectví, lékařství, automatizaci, chemickém průmyslu a ekonometrii.

U dynamických systémů se obvykle obtížně posuzuje odhad z hlediska jeho konvergence ke skutečné hodnotě, a proto se častěji hodnotí odhad z hlediska velikosti jeho chyby. V některých případech může estimátor produkovat neakceptovatelně velké chyby odhadu, což se někdy označuje jako divergence odhadu. To se také často stávalo v prvních pokusech o reálné nasazení Kalmanova filtru v avionice. Vědecká veřejnost byla výsledky natolik zklamána, že se o Kalmanově filtru hovořilo jako o „nejhorším vynálezu dekády“.

Z výše uvedeného je zřejmé, že stejně důležité jako návrh řešení, je ověření jeho korektnosti a robustnosti, tedy vyhodnocení kvality odhadu. Většina estimátorů je navržena optimalizací nějakého kritéria, proto by bylo přirozené hodnotit kvalitu odhadu právě podle tohoto návrhového kritéria. Situace ovšem není tak jednoznačná. Při výběru návrhového kritéria je třeba mít na paměti matematickou řešitelnost celého pro-

blému, proto se v drtivé většině případů volí jako kritérium střední kvadratická chyba (*Mean Squared Error*, MSE). Kritérium MSE má ideální návrhové vlastnosti (lineární derivace), ale má i své nevýhody jako je přílišná pesimističnost - protežování velkých chyb na úkor malých. Proto, ačkoliv to zní paradoxně, je vhodné hodnotit kvalitu odhadu nejen pomocí návrhového kritéria, ale i jinými kritérii, která prověří jiné složky a módy kvality odhadu.

Hodnocení kvality odhadu je mezi širokou vědeckou veřejností neprávem přehlíženou problematikou. Ačkoliv se jedná o velmi důležité téma, není známa žádná doporučená metodika přístupu k tomuto problému a práci na toto téma publikuje pouze několik málo autorů. První, kdo se tomuto tématu začal systematicky věnovat, byl v 80. letech minulého století Y. Bar-Shalom. Ještě produktivnější je Bar-Shalomův bývalý doktorand X. Rong Li, který je de-facto spoluautorem většiny prací na toto téma. O tom, že se jedná o téma, které je v poslední době na vzestupu, také svědčí, že většina podkladů pro tuto práci je datována do současného milénia.

Existuje řada metrik hodnotících chybu odhadu stavu, některé uvažují pouze samotný bodový odhad, některé navíc respektují kovarianční matici chyby odhadu poskytovanou estimátorem. Metriky chyby odhadu mohou zkoumat buď absolutní nebo relativní chybu odhadu, anebo míru četnosti vyhovujících, resp. nevyhovujících odhadů. Metriky mohou být označeny za optimistické, vyvážené či pesimistické na základě jejich vztahu k potlačování, resp. protežování velkých chyb. Mnohdy je po analýze chyby odhadu také důležité vyhodnotit robustnost či věrohodnost odhadu. K tomuto účelu slouží metriky kovarianční a MSE matice. Tyto metody většinou nějakým způsobem hodnotí tendenci estimátorů k nadhodnocování či podhodnocování kvality jejich vlastních odhadů.

Cílem práce je poskytnout přehled možností pro hodnocení kvality bodových odhadů stavu, ilustrovat typické využití hodnocení kvality a doporučit metodiku hodnocení kvality.

Práce je rozčleněna následujícím způsobem. Druhá kapitola představuje úvod do úlohy odhadu stavu, formulaci a řešení problému. Ve třetí kapitole jsou podrobně specifikovány cíle diplomové práce. Čtvrtá kapitola uvádí metriky hodnocení kvality odhadu stavu. V páté kapitole jsou představeny metriky kvality hodnotící věrohodnost a limitní poznatelnost odhadů. Šestá kapitola shrnuje předchozí poznatky a dává návod, jak přistupovat metodicky k hodnocení kvality odhadu. Sedmá kapitola ukazuje aplikaci hodnocení kvality na tři typické problémy z různých oborů. Konečně osmá, závěrečná kapitola shrnuje výsledky práce.

Kapitola 2

Úloha odhadu stavu

V části 2.1 této kapitoly bude zaměřena pozornost na popis dynamického stochastického systému diskrétního v čase. Část 2.2 představí formulaci úlohy odhadu stavu tohoto systému a ukáže obecné řešení úlohy. Část 2.3 stručně shrne myšlenky globálních nelineárních metod řešení úlohy odhadu stavu. Konečně poslední část 2.4 se bude věnovat lokálním nelineárním metodám odhadu stavu poskytujícím bodové odhady, jejichž kvalita bude hodnocena v dalších kapitolách.

2.1 Popis systému

V této práci bude uvažován nelineární stochastický systém diskrétní v čase popsany vztahy:

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k) + \mathbf{w}_k, \quad k = 0, 1, 2, \dots, \quad (2.1a)$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k, \quad k = 0, 1, 2, \dots, \quad (2.1b)$$

kde vektory $\mathbf{x}_k \in \mathbb{R}^{n_x}$ a $\mathbf{z}_k \in \mathbb{R}^{n_z}$ reprezentují stav systému a měření v čase k . Vektory $\mathbf{w}_k \in \mathbb{R}^{n_x}$ a $\mathbf{v}_k \in \mathbb{R}^{n_z}$ představují stavový šum a šum měření, jejichž hustoty pravděpodobnosti $p(\mathbf{w}_k)$, resp. $p(\mathbf{v}_k)$ jsou známé. Pokud nebude uvedeno jinak, budou tyto šумы v celé práci uvažovány jako gaussovské s nulovou střední hodnotou a kovariančními maticemi \mathbf{Q}_k , resp. \mathbf{R}_k . Oba šумы jsou bílé, vzájemně nezávislé, a také nezávislé na počátečním stavu \mathbf{x}_0 popsaném hustotou pravděpodobnosti $p(\mathbf{x}_0)$, která je v této práci opět uvažována jako gaussovská se střední hodnotou $\hat{\mathbf{x}}_0$ a kovarianční maticí \mathbf{P}_0 . Taktéž známé jsou vektorové funkce $\mathbf{f}_k : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_x}$ a $\mathbf{h}_k : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{n_z}$.

Stavový šum a šum měření jsou považovány za nezávislé bez újmy na obecnosti, neboť lze jednoduchou transformací převést korelované šумы na nezávislé [34]. Stejně tak bylo dokázáno, že aditivita šumů není na úkor obecnosti [28].

Stav systému \mathbf{x}_k vyjadřuje veškerou informaci o dynamice systému, tím pádem je to minimální množina stavových proměnných potřebných pro úplný popis chování dynamického systému. Stav \mathbf{x}_k není přímo měřitelný, ale pouze nepřímo prostřednictvím měření \mathbf{z}_k . Systém popsany rovnicemi (2.1) může být alternativně vyjádřen pomocí podmíněných hustot pravděpodobnosti:

$$p(\mathbf{x}_{k+1} | \mathbf{x}_k), \quad k = 0, 1, 2, \dots, \quad (2.2a)$$

$$p(\mathbf{z}_k | \mathbf{x}_k), \quad k = 0, 1, 2, \dots, \quad (2.2b)$$

kde (2.2a) je přechodová hustota pravděpodobnosti stavu a (2.2b) podmíněná hustota pravděpodobnosti měření.

Důsledkem definice stavu a vlastností stavového šumu je, že $\{\mathbf{x}_k\}_{k=0}^{\infty}$ je markovský proces, tedy platí:

$$p(\mathbf{x}_{k+1}|\mathbf{x}_k) = p(\mathbf{x}_{k+1}|\mathbf{x}_k, \mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots). \quad (2.3)$$

Tato vlastnost říká, že současný stav v sobě akumuluje veškerou informaci o minulosti. Poznamenejme, že podle [13] je dokonce markovost příbuzná s kauzalitou.

2.2 Formulace a řešení úlohy odhadu

Cílem odhadu stavu stochastických systémů je na základě měřených dat a jisté apriorní informace o sledovaném systému navrhnout co nejlepší odhad stavu \mathbf{x}_k .

Odhad stavu \mathbf{x}_k je na základě znalosti měření až do času l reprezentován aposteriorní hustotou $p(\mathbf{x}_k|\mathbf{z}^l)$, kde \mathbf{z}^l je vektor měření až do času l , tedy $\mathbf{z}^l \triangleq [\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_l]^T$.

S ohledem na vztah mezi k a l dělíme úlohu odhadu stavu na:

- úlohu predikce pro $k > l$,
- úlohu filtrace pro $k = l$,
- úlohu vyhlazování¹ pro $k < l$.

Hustota pravděpodobnosti $p(\mathbf{x}_k|\mathbf{z}^l)$ se získá řešením tzv. funkcionálních rekurzivních vztahů (FRV). Obecné řešení úlohy filtrace je dáno vztahem

$$p(\mathbf{x}_k|\mathbf{z}^k) = \frac{p(\mathbf{x}_k|\mathbf{z}^{k-1}) p(\mathbf{z}_k|\mathbf{x}_k)}{\int p(\mathbf{x}_k|\mathbf{z}^{k-1}) p(\mathbf{z}_k|\mathbf{x}_k) d\mathbf{x}_k}, \quad (2.4)$$

kde jedнокroková prediktivní hustota $p(\mathbf{x}_k|\mathbf{z}^{k-1})$ je dána vztahem

$$p(\mathbf{x}_k|\mathbf{z}^{k-1}) = \int p(\mathbf{x}_{k-1}|\mathbf{z}^{k-1}) p(\mathbf{x}_k|\mathbf{x}_{k-1}) d\mathbf{x}_{k-1}. \quad (2.5)$$

Hustota pravděpodobnosti vícekové predikce $p(\mathbf{x}_k|\mathbf{z}^l)$ pro $k > l$ je určena rekurzivním vztahem

$$p(\mathbf{x}_k|\mathbf{z}^l) = \int p(\mathbf{x}_{k-1}|\mathbf{z}^l) p(\mathbf{x}_k|\mathbf{x}_{k-1}) d\mathbf{x}_{k-1}. \quad (2.6)$$

Výpočet vícekové vyhlazovací hustoty pravděpodobnosti $p(\mathbf{x}_k|\mathbf{z}^l)$ pro $k < l$ popisuje vztah

$$p(\mathbf{x}_k|\mathbf{z}^l) = p(\mathbf{x}_k|\mathbf{z}^k) \int \frac{p(\mathbf{x}_{k+1}|\mathbf{z}^l)}{p(\mathbf{x}_{k+1}|\mathbf{z}^k)} p(\mathbf{x}_{k+1}|\mathbf{x}_k) d\mathbf{x}_{k+1}. \quad (2.7)$$

¹Ize se setkat i s termíny retrodikce či interpolace, avšak termín vyhlazování lépe vystihuje podstatu slova „smoothing“ zavedeného v anglicky psané literatuře [9].

Počáteční podmínka pro rekurzi je vyjádřena hustotou pravděpodobnosti počátečního stavu

$$p(\mathbf{x}_0|\mathbf{z}^{-1}) = p(\mathbf{x}_0). \quad (2.8)$$

Vztahy (2.4) a (2.5) řešící úlohu filtrace jsou natolik významné a často používané, že se o nich hovoří jako o Bayesových rekurzivních vztazích (BRV). Tato práce se bude věnovat hlavně úloze filtrace, proto se dále hovoří pouze o BRV, a nikoliv o FRV.

Exaktní řešení Bayesových rekurzivních vztahů je ale známé pouze pro několik speciálních případů. Nejvýznamnější z nich je lineární gaussovský systém, jehož řešením jsou vztahy známého Kalmana filtru. V ostatních případech se musíme dopustit jistých nepřesností, tedy provádět aproximace.

2.3 Globální metody odhadu stavu

Metody založené na BRV se nazývají globální nelineární filtrační metody. Díky tomu, že se tyto metody pokouší najít podmíněné hustoty pravděpodobnosti nebo jejich aproximace, poskytují výsledek platný v téměř celém stavovém prostoru, z čehož vyplývá jejich název. Zásadní nevýhodou těchto metod je v drtivé většině případů vysoká výpočetní náročnost a složitý matematický aparát nutný k jejich návrhu.

V zásadě existují tři přístupy k návrhu globálních metod řešící BRV:

- Analytický přístup aproximuje filtrační hustotu, prediktivní hustotu, stavový šum, šum měření a počáteční stav pomocí směsi gaussovských rozdělání. Příkladem je filtr gaussovských směsí (*Gaussian Sum Filter*) [32].
- Numerický přístup řeší integrály v BRV nahrazením spojitého stavového prostoru sítí konečného počtu obvykle ortogonálních a ekvidistantních bodů. Příkladem je filtr bodových mas (*Point-Mass Filter*) [18].
- Simulační přístup využívá aproximace filtrační hustoty velkým množstvím náhodných vzorků spolu s jejich váhovým ohodnocením. Příkladem je částicový filtr (*Particle Filter / Sequential Monte Carlo Method*) [8].

2.4 Lokální metody odhadu stavu

V některých případech postačuje znalost bodového odhadu stavu \mathbf{x}_k a není nutné znát celou hustotu pravděpodobnosti $p(\mathbf{x}_k|\mathbf{z}^k)$. Protože je odhad funkcí předchozích měření, označme jej $\hat{\mathbf{x}}_k(\mathbf{z}^k)$. Metodám poskytujícím bodové odhady stavu se říká lokální metody, neboť poskytují odhad platný pouze v malém okolí bodu $\hat{\mathbf{x}}_k(\mathbf{z}^k)$. Tyto metody často aproximují podmíněnou hustotu pravděpodobnosti stavu gaussovským rozděláním.

K získání bodového odhadu $\hat{\mathbf{x}}_k(\mathbf{z}^k)$ je třeba zvolit nejprve optimalizační kritérium. Z praktických důvodů se nejčastěji volí střední kvadratická chyba (*Mean Squared Error*, MSE), jejíž derivace je lineární a usnadňuje další řešení. Výsledkem je

$$\hat{\mathbf{x}}_k^*(\mathbf{z}^k) = \arg \min_{\hat{\mathbf{x}}_k(\mathbf{z}^k)} \mathbb{E} \left[(\mathbf{x}_k - \hat{\mathbf{x}}_k(\mathbf{z}^k))^T (\mathbf{x}_k - \hat{\mathbf{x}}_k(\mathbf{z}^k)) | \mathbf{z}^k \right], \quad (2.9)$$

$$\hat{\mathbf{x}}_k^*(\mathbf{z}^k) = \mathbb{E} [\mathbf{x}_k | \mathbf{z}^k], \quad (2.10)$$

tedy bodový odhad, který je optimální ve smyslu střední kvadratické chyby je dán podmíněnou střední hodnotou (2.10).

Než přistoupíme k detailnímu popisu estimátorů, které budou použity v této práci, podíváme se nejprve na tzv. gaussovský filtr (*Gaussian Filter*, GF). Tento filtr byl poprvé odvozen v [12] a je v podstatě zobecněním většiny lokálních filtrů, i když byly neřídko tyto filtry odvozeny za jiných předpokladů. GF je odvozen za předpokladu, že podmíněná hustota stavu je gaussovská. Následuje odvození vztahů definujících GF.

Začneme prediktivním krokem, ve kterém aproximujeme jedнокrokovou prediktivní hustotu $p(\mathbf{x}_k | \mathbf{z}^{k-1})$ gaussovským rozdělením

$$p(\mathbf{x}_k | \mathbf{z}^{k-1}) \approx \mathcal{N}\{\mathbf{x}_{k|k-1}; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}\}, \quad (2.11)$$

kde se momenty rozdělení vypočítají využitím (2.1a) následovně:

$$\hat{\mathbf{x}}_{k|k-1} \triangleq \mathbb{E}[\mathbf{x}_k | \mathbf{z}^{k-1}] = \mathbb{E}[\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) | \mathbf{z}^{k-1}], \quad (2.12)$$

$$\mathbf{P}_{k|k-1} = \mathbb{E}[(\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \hat{\mathbf{x}}_{k|k-1})(\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \hat{\mathbf{x}}_{k|k-1})^T | \mathbf{z}^{k-1}] + \mathbf{Q}_{k-1}. \quad (2.13)$$

Vztahy (2.12) a (2.13) mohou být přepsány do formy integrálů:

$$\hat{\mathbf{x}}_{k|k-1} = \int \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}^{k-1}) d\mathbf{x}_{k-1}, \quad (2.14)$$

$$\mathbf{P}_{k|k-1} = \int [\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \hat{\mathbf{x}}_{k|k-1}] [\mathbf{f}_{k-1}(\mathbf{x}_{k-1}) - \hat{\mathbf{x}}_{k|k-1}]^T p(\mathbf{x}_{k-1} | \mathbf{z}^{k-1}) d\mathbf{x}_{k-1} + \mathbf{Q}_{k-1}, \quad (2.15)$$

kde filtrační hustota pravděpodobnosti

$$p(\mathbf{x}_{k-1} | \mathbf{z}^{k-1}) = \mathcal{N}\{\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}\} \quad (2.16)$$

byla aproximována gaussovským rozdělením.

Pro odvození filtračního kroku předpokládáme, že $[\mathbf{x}_k, \mathbf{h}_k(\mathbf{x}_k)]$ má sdružené gaussovské rozdělení podmíněné \mathbf{z}^{k-1} . Potom jsou střední hodnota a kovarianční matice filtrační hustoty pravděpodobnosti dány vztahy (viz [36])

$$\hat{\mathbf{x}}_{k|k} \triangleq \mathbb{E}[\mathbf{x}_k | \mathbf{z}^k] = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}), \quad (2.17)$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{P}_{k|k-1}^{zz} \mathbf{K}_k^T, \quad (2.18)$$

kde

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}^{xz} (\mathbf{P}_{k|k-1}^{zz})^{-1} \quad (2.19)$$

je tzv. Kalmanův zisk a prediktivní kovarianční matice $\mathbf{P}_{k|k-1}^{xz}$ je dána

$$\begin{aligned} \mathbf{P}_{k|k-1}^{xz} &= \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})^T | \mathbf{z}^{k-1}] \\ &= \int (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{h}_k(\mathbf{x}_k) - \hat{\mathbf{z}}_{k|k-1})^T p(\mathbf{x}_k | \mathbf{z}^{k-1}) d\mathbf{x}_k. \end{aligned} \quad (2.20)$$

Vztahy pro výpočet predikce měření a příslušné kovarianční matice jsou

$$\hat{\mathbf{z}}_{k|k-1} = \int \mathbf{h}_k(\mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}^{k-1}) d\mathbf{x}_k, \quad (2.21)$$

$$\begin{aligned}
\mathbf{P}_{k|k-1}^{zz} &= E[(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})^T | \mathbf{z}^{k-1}] \\
&= E[(\mathbf{h}_k(\mathbf{x}_k) - \hat{\mathbf{z}}_{k|k-1}) \times (\mathbf{h}_k(\mathbf{x}_k) - \hat{\mathbf{z}}_{k|k-1})^T | \mathbf{z}^{k-1}] + \mathbf{R}_k \\
&= \int (\mathbf{h}_k(\mathbf{x}_k) - \hat{\mathbf{z}}_{k|k-1})(\mathbf{h}_k(\mathbf{x}_k) - \hat{\mathbf{z}}_{k|k-1})^T p(\mathbf{x}_k | \mathbf{z}^{k-1}) d\mathbf{x}_k + \mathbf{R}_k,
\end{aligned} \tag{2.22}$$

kde jednokroková prediktivní hustota pravděpodobnosti

$$p(\mathbf{x}_k | \mathbf{z}^{k-1}) = \mathcal{N}\{\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}\} \tag{2.23}$$

byla aproximována gaussovským rozdělením.

Vztahy (2.11)-(2.23) představují gaussovský filtr. Pro přehlednost bude uveden kompletní algoritmus:

Algoritmus: Gaussovský filtr

Krok 1: (*inicializace*) Necht' současný čas je $k = 0$ a apriorní počáteční podmínka je definována jejími prvními dvěma momenty

$$\hat{\mathbf{x}}_{0|-1} \triangleq E[\mathbf{x}_0] = \hat{\mathbf{x}}_0, \tag{2.24}$$

$$\mathbf{P}_{0|-1} \triangleq E[(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|-1})(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|-1})^T] = \mathbf{P}_0. \tag{2.25}$$

Krok 2: (*filtrace*) Filtrační odhad $\hat{\mathbf{x}}_{k|k}$ a filtrační kovarianční matice $\mathbf{P}_{k|k}$ se vypočtou pomocí rekurzivními vztahy

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}), \tag{2.26}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{P}_{k|k-1}^{zz} \mathbf{K}_k^T, \tag{2.27}$$

kde

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}^{xz} (\mathbf{P}_{k|k-1}^{zz})^{-1} \tag{2.28}$$

je filtrační zisk a odhad měření $\hat{\mathbf{z}}_{k|k-1}$ je dán střední hodnotou

$$\hat{\mathbf{z}}_{k|k-1} = E[\mathbf{z}_k | \mathbf{z}^{k-1}]. \tag{2.29}$$

Prediktivní kovarianční matice $\mathbf{P}_{k|k-1}^{xz}$ a $\mathbf{P}_{k|k-1}^{zz}$ jsou vypočteny

$$\begin{aligned}
\mathbf{P}_{k|k-1}^{zz} &= E[(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})^T | \mathbf{z}^{k-1}] \\
&= E[(\mathbf{h}_k(\mathbf{x}_k) - \hat{\mathbf{z}}_{k|k-1})(\mathbf{h}_k(\mathbf{x}_k) - \hat{\mathbf{z}}_{k|k-1})^T | \mathbf{z}^{k-1}] + \mathbf{R}_k,
\end{aligned} \tag{2.30}$$

$$\mathbf{P}_{k|k-1}^{xz} = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})^T | \mathbf{z}^{k-1}]. \tag{2.31}$$

Krok 3: (*predikce*) Prediktivní odhad $\hat{\mathbf{x}}_{k+1|k}$ a prediktivní kovarianční matice $\mathbf{P}_{k+1|k}$ jsou dány

$$\hat{\mathbf{x}}_{k+1|k} = E[\mathbf{x}_{k+1} | \mathbf{z}^k] = E[\mathbf{f}_k(\mathbf{x}_k) | \mathbf{z}^k], \tag{2.32}$$

$$\mathbf{P}_{k+1|k} = E[(\mathbf{f}_k(\mathbf{x}_k) - \hat{\mathbf{x}}_{k+1|k})(\mathbf{f}_k(\mathbf{x}_k) - \hat{\mathbf{x}}_{k+1|k})^T | \mathbf{z}^k] + \mathbf{Q}_k. \tag{2.33}$$

Necht' $k = k + 1$. Algoritmus pak pokračuje **Krokem 2**.

Ačkoliv má gaussovský filtr lineární strukturu, vyžaduje výpočet integrálů (2.14, 2.15, 2.20, 2.21, 2.22), což lze provést analyticky pouze v několika speciálních případech. Významným speciálním případem je situace, kdy jsou vektorové funkce \mathbf{f}_k a \mathbf{h}_k lineární. V tomto případě bychom se dostali ke vztahům definujícím Kalmanův filtr (KF). Obecně jsou tyto vektorové funkce nelineární. Z tohoto důvodu je třeba přistoupit k další aproximaci, abychom byli schopni tyto integrály vypočítat. Možnosti jsou v zásadě dvě.

Historicky starší přístup je aproximovat nelineární funkce \mathbf{f}_k a \mathbf{h}_k v modelu systému (2.1). Toho lze dosáhnout například linearizací těchto funkcí pomocí Taylorova rozvoje prvního, resp. druhého řádu, což vede na rozšířený Kalmanův filtr (*Extended Kalman Filter*, EKF), iterační filtr, resp. filtr druhého řádu [1, 34]. Lze také využít Stirlingovu polynomiální interpolaci, která může být chápána jako Taylorův rozvoj, kde jsou derivace nahrazeny diferencemi. Tento přístup vede na diferenční lokální filtry (*Divided Difference Filters*) [28]. Nebo lze využít rozvoje pomocí ortogonálních tzv. Fourier-Hermitových řad. Příkladem je Fourier-Hermitův KF [31].

Mladší jsou filtry využívající transformaci charakteristických bodů, tzv. σ -bodů. Jde o numerické řešení integrálů využitím množiny deterministicky zvolených vážených bodů, přičemž popis systému je ponechán beze změn. Mezi zástupce těchto filtrů patří unscentovaný KF (*Unscented Kalman Filter*, UKF) [14], kvadrurní KF [3], kubaturní KF [2] nebo jejich zobecnění - stochastický integrační filtr (*Stochastic Integration Filter*) [10].

V této práci budou použity dva filtry. V praxi stále nejpoužívanější rozšířený Kalmanův filtr a moderní bezderivační unscentovaný Kalmanův filtr.

2.4.1 Rozšířený Kalmanův filtr

Mezi lety 1959-1961 publikoval R. E. Kalman články o tehdy novém pohledu na filtraci. Navrhl přejít od frekvenčního popisu ke stavovému, a později odvodil vztahy definující Kalmanův filtr. Tento filtr má z praktického hlediska několik omezujících předpokladů, přičemž tím nejzávažnějším je, že systém musí být lineární. Přitom však drtivá většina praktických problémů je nelineární. V této době se také naplno rozjížděly vesmírné závody a řešení nelineárního problému filtrace bylo klíčové pro dosažení cíle. Prvním, a v současné navigaci a lokalizaci stále ještě preferovaným řešením, bylo linearizovat vektorové funkce \mathbf{f}_k a \mathbf{h}_k pomocí Taylorova rozvoje.

Abychom mohli použít Taylorův rozvoj, musí být zaručena diferencovatelnost funkcí \mathbf{f}_k a \mathbf{h}_k . Pak pro Taylorův rozvoj prvního řádu funkce \mathbf{f}_k v bodě $\mathbf{x}_k = \hat{\mathbf{x}}_k$ platí

$$\mathbf{f}_k(\mathbf{x}_k) \approx \mathbf{f}_k(\hat{\mathbf{x}}_k) + \mathbf{f}'_k(\hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k). \quad (2.34)$$

Rozšířený Kalmanův filtr aproximuje vektorové funkce \mathbf{f}_k a \mathbf{h}_k Taylorovým rozvojem prvního řádu v bodě $\mathbf{x}_k = \hat{\mathbf{x}}_{k|k}$ resp. $\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}$ a členy vyšších řádů jsou zanedbány. Tím dostaneme do integrálů (2.14, 2.15, 2.20, 2.21, 2.22) lineární funkce a tyto integrály

se stávají analyticky řešitelné. Řešením těchto integrálů dostaneme

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{f}_k(\hat{\mathbf{x}}_{k|k}), \quad (2.35)$$

$$\mathbf{P}_{k+1|k} = \mathbf{F}_k(\hat{\mathbf{x}}_{k|k})\mathbf{P}_{k|k}\mathbf{F}_k^T(\hat{\mathbf{x}}_{k|k}) + \mathbf{Q}_k \quad (2.36)$$

$$\mathbf{P}_{k|k-1}^{xz} = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\hat{\mathbf{x}}_{k|k-1}) \quad (2.37)$$

$$\hat{\mathbf{z}}_{k|k-1} = \mathbf{h}_k(\hat{\mathbf{x}}_{k|k-1}) \quad (2.38)$$

$$\mathbf{P}_{k|k-1}^{zz} = \mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1})\mathbf{P}_{k|k-1}\mathbf{H}_k^T(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{R}_k \quad (2.39)$$

kde

$$\mathbf{H}_k(\hat{\mathbf{x}}_{k|k-1}) = \left. \frac{\partial \mathbf{h}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k} \right|_{\hat{\mathbf{x}}_{k|k-1}} \quad (2.40)$$

je matice n_z/n_x . kde

$$\mathbf{F}_k(\hat{\mathbf{x}}_{k|k}) = \left. \frac{\partial \mathbf{f}_k(\mathbf{x}_k)}{\partial \mathbf{x}_k} \right|_{\hat{\mathbf{x}}_{k|k}} \quad (2.41)$$

je matice n_x/n_x .

Pro úplnost dodejme, že musíme znát počáteční podmínku $\hat{\mathbf{x}}_{0|-1}$ a $\mathbf{P}_{0|-1}$. Pak už se jen střídá filtrační a prediktivní krok. Algoritmus EKF je formálně shodný s algoritmem GF na straně 11. Důležité je uvědomit si několik rozdílů mezi Kalmanovým filtrem a rozšířeným Kalmanovým filtrem. Za prvé kovarianční matice $\mathbf{P}_{k|k-1}$ a $\mathbf{P}_{k|k}$ jsou funkcemi měření, a nelze je proto předpočítat dopředu jako v případě KF. A hlavně, EKF poskytuje pouze aproximace skutečných hustot na rozdíl od exaktního KF, jelikož je aproximace prováděna pouze v jednom konkrétním bodě stavového prostoru, proto je platnost výsledků pouze lokální a tudíž může filtr divergovat [29].

2.4.2 Unscentovaný Kalmanův filtr

Ačkoliv je EKF pravděpodobně nejpoužívanější estimační algoritmus, jeho více než 45 let použití v praxi odhalilo jeho vážné nedostatky. Je poměrně těžké tento algoritmus implementovat a následně vyladit [40]. A pro dostatečnou spolehlivost algoritmu musí být systém blízký lineárnímu. Tyto slabiny vycházejí z aproximace linearizací. V roce 1995 byl zveřejněn nový přístup k filtraci nelineárních systémů založený na transformaci deterministicky zvolených bodů nelineární funkcí. Této transformaci bylo později zvoleno jméno unscentovaná transformace (UT). Tato transformace a přístup si prošly velmi bouřlivým a intenzivním vývojem, až byli tvůrci UT S. J. Julier a J. K. Uhlmann v roce 2004 vyzváni k sepsání souhrnného článku [14], který se stal nejcitovanějším článkem poslední dekády v oblasti automatického řízení.

Unscentovaná transformace slouží k aproximativnímu výpočtu střední hodnoty, kovariance a vzájemné kovariance transformované náhodné veličiny $\mathbf{y} = \mathbf{g}(\mathbf{x})$, za předpokladu známé nelineární funkce \mathbf{g} , známé střední hodnoty $\hat{\mathbf{x}} = \mathbf{E}[\mathbf{x}]$ a kovarianční matice $\mathbf{P} = \text{var}[\mathbf{x}]$ náhodné proměnné \mathbf{x} . UT k tomu využívá množinu σ -bodů $\{\mathcal{X}_i\}_{i=0}^{2n_x}$

s odpovídajícími váhami $\{\mathcal{X}_i\}_{i=0}^{2n_x}$ takto:

$$\begin{aligned}\mathcal{X}_0 &= \hat{\mathbf{x}}, & \mathcal{W}_0 &= \frac{\kappa}{n_x + \kappa}, \\ \mathcal{X}_j &= \hat{\mathbf{x}} + \sqrt{(n_x + \kappa)} \left(\sqrt{\mathbf{P}^{xx}} \right)_j, & \mathcal{W}_j &= \frac{1}{2(n_x + \kappa)}, \\ \mathcal{X}_{j+1} &= \hat{\mathbf{x}} - \sqrt{(n_x + \kappa)} \left(\sqrt{\mathbf{P}^{xx}} \right)_j, & \mathcal{W}_{j+1} &= \frac{1}{2(n_x + \kappa)},\end{aligned}\quad (2.42)$$

kde $j = 1, \dots, n_x$, výraz $\left(\sqrt{\mathbf{P}^{xx}} \right)_j$ znamená j -tý sloupec matice $\sqrt{\mathbf{P}^{xx}}$, což je maticová dekompozice \mathbf{P}^{xx} tak, že platí $\sqrt{\mathbf{P}^{xx}} \left[\sqrt{\mathbf{P}^{xx}} \right]^T = \mathbf{P}^{xx}$ a proměnná κ je parametr změny měřítka (*scaling*). Tato dekompozice může být vypočtena například Choleského dekompozicí nebo dekompozicí na singulární čísla (SVD). Takto získaná množina σ -bodů má přinejmenším první dva momenty stejné jako náhodná proměnná \mathbf{x} , tj.

$$\hat{\mathbf{x}} = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{X}_i, \quad (2.43)$$

$$\mathbf{P}^{xx} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{X}_i - \hat{\mathbf{x}}) (\mathcal{X}_i - \hat{\mathbf{x}})^T. \quad (2.44)$$

Každý σ -bod transformujeme nelineární funkcí

$$\mathcal{Y}_i = \mathbf{g}(\mathcal{X}_i), \forall i. \quad (2.45)$$

a vypočteme aproximativní charakteristiky

$$\hat{\mathbf{y}} = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{Y}_i, \quad (2.46)$$

$$\mathbf{P}^{yy} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{Y}_i - \hat{\mathbf{y}}) (\mathcal{Y}_i - \hat{\mathbf{y}})^T, \quad (2.47)$$

$$\mathbf{P}^{xy} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{X}_i - \hat{\mathbf{x}}) (\mathcal{Y}_i - \hat{\mathbf{y}})^T. \quad (2.48)$$

Algoritmus UKF je opět formálně shodný s algoritmem GF na straně 11, kde prediktivní momenty měření $\hat{\mathbf{z}}_{k|k-1}$, $\mathbf{P}_{k|k-1}^{zz}$, $\mathbf{P}_{k|k-1}^{xz}$ a prediktivní momenty stavu $\hat{\mathbf{x}}_{k+1|k}$, $\mathbf{P}_{k+1|k}$ jsou vypočteny pomocí UT (2.42-2.48).

Kapitola 3

Cíle práce

Každý estimátor je navržen na základě jistých více či méně restriktivních předpokladů. Ve většině případů (např. díky nelinearitě systému) není ale možné platnost takových předpokladů ověřit. Systém je z různých důvodů nemusí splňovat. Dá se říci, že kvalita odhadu je závislá na datech. Proto je vždy dobré kvalitu odhadu (poskytovaného estimátorem pro daný systém) nějakým statistickým způsobem vyhodnotit. Toto se nejčastěji provádí simulací Monte-Carlo.

I když jsou požadavky estimátorů na systém beze zbytku splněny, jakmile je systém nelineární nebo negaussovský, nelze o kvalitě odhadu říci apriori vůbec nic. V těchto případech je odhad vždy výsledkem nějakých aproximací provedených v rámci odvození estimátoru, a tak není možné určit jejich kumulativní efekt na výsledek.

Ať již nastane situace z prvního či druhého odstavce, jsou výsledky platné pouze pro daný konkrétní scénář, tj. pro dané funkce \mathbf{f} a \mathbf{h} , vlastnosti šumů a počáteční podmínky).

Výhodou počítačových simulací na rozdíl od reality je znalost skutečného stavu. Proto lze hodnotit kvalitu bodového odhadu poskytovaného lokálními filtry poměrně přirozeně, například srovnáním skutečného stavu přímo s jeho bodovým odhadem. Hodnocení kvality odhadu globálních filtrů už tak přímočaré není, neboť většinou není k dispozici filtrační hustota, kterou se snaží tyto filtry nalézt (s výjimkou speciálních případů). Proto se tato práce soustředí na hodnocení kvality odhadů poskytovaných lokálními filtry.

Jak již bylo řečeno, lokální filtry jsou odvozeny optimalizací určitého návrhového kritéria. Kritéria (nejčastěji MSE) nejsou volena jen podle vhodnosti pro daný problém, ale především podle toho, zda vůbec umožňují dopracovat se k řešení úlohy. Kritérium v různých aplikacích nemusí být žádoucí pro hodnocení kvality odhadu a požaduje se hodnocení i podle jiných kritérií (např. absolutní chyba). Tato kritéria mohou mít z jistého pohledu vhodnější vlastnosti než pesimistické MSE, které příliš penalizuje velké chyby nebo mohou vedle filtrační střední hodnoty hodnotit i kovarianční matici chyby odhadu.

Cílem práce je:

- představit různá kritéria hodnotící kvalitu odhadu ze všech možných úhlů pohledu,
- vybědnout k užití méně známých, přesto zajímavých a vypovídajících hodnocení,

- rozšířit povědomí o tématu, které je v české a dokonce i světové literatuře neprávem opomíjeno,
- předvést nasazení těchto hodnocení ve třech konkrétních typických problémech z různých odvětví,
- doporučit, jak se v tomto tématu zorientovat a která hodnocení preferovat.

Kapitola 4

Metriky chyby odhadu

V této kapitole budou předvedeny metriky¹ vhodné k ohodnocení kvality odhadů z hlediska jejich absolutní a relativní chyby. Kritéria kvality hodnotící kovarianční matici $\mathbf{P}_{k|k}$ dodanou estimátorem jsou předmětem kapitoly 5.

Pro zjednodušení zavedení kritérií kvality bude v následujících kapitolách užívána tato konvence. Estimátory dodávají odhad $\hat{\mathbf{x}}_{k|k}$ skutečného stavu \mathbf{x}_k pro okamžik k za podmínky pozorování do okamžiku k . Veškeré parametry simulovaného systému jsou známy, včetně skutečných stavů \mathbf{x}_k , díky čemuž lze pracovat s chybou odhadu $\tilde{\mathbf{x}}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$. Pokud bude v rámci přehlednosti vynechán index časových kroků k , bude automaticky uvažováno, že se jedná o danou proměnou právě v čase k . Pro každý časový okamžik k bude k dispozici M nezávislých iterací simulace Monte Carlo (MC), a tak přibude informace, že se jedná o i -tou iteraci simulace MC. Značení $\hat{\mathbf{x}}_k(i)$ tedy bude značit odhad stavu $\hat{\mathbf{x}}_{k|k}$ dodaný estimátorem z i -té iterace simulace MC.

Metriky mohou být klasifikovány dvěma způsoby:

- podle druhu výsledku, který poskytují na metriky absolutní chyby, relativní chyby a míry četnosti (mimořádných událostí),
- podle tolerance chyb na metriky optimistické (jak dobrý odhad je, tolerance velkých chyb), pesimistické (jak špatný odhad je, penalizace velkých chyb) a vyvážené (ani optimistické ani pesimistické).

Kapitola představí jednotlivé metriky, a to v pořadí prvního uvedeného dělení, přičemž vždy bude řečeno, jestli je daná metrika optimistická, pesimistická či vyvážená, jaké má využití a jaké je její typické nasazení. Na závěr se bude kapitola věnovat problému ovlivňujícím všechny metriky, problému fyzikálně různorodých stavů a stavů řádově odlišných hodnot.

4.1 Metriky absolutní chyby

Metriky absolutní chyby jsou zdaleka nejpopulárnějším a nejpřirozenějším měřítkem kvality odhadu. Je důležité si uvědomit, že vztahy zde představené jsou ve skutečnosti pouze odhadem, který se snaží aproximovat skutečné hodnoty metriky z konečného

¹Termín metrika je chápán volně jako hodnota oceňující kvalitu odhadu, nikoliv v striktně matematickém pojetí. V anglicky psané odborné literatuře se termín „metric“ běžně pro tento účel používá.

počtu realizací. Když tedy bude řeč například o mediánu chyby, půjde ve skutečnosti pouze o odhad mediánu z M vzorků.

4.1.1 Odmocnina ze střední kvadratické chyby - RMSE

Nejznámější a nejrozšířenější metrikou je odmocnina ze střední kvadratické chyby (*Root Mean Squared Error*) definovaná vztahem

$$\text{RMSE}(\hat{\mathbf{x}}_k) = \sqrt{\frac{1}{M} \sum_{i=1}^M \|\tilde{\mathbf{x}}_k(i)\|^2}, \quad (4.1)$$

kde $\|\tilde{\mathbf{x}}_k(i)\| = \sqrt{\tilde{\mathbf{x}}_k^T(i)\tilde{\mathbf{x}}_k(i)}$ je eukleidovská norma chyby odhadu.

RMSE je nejpřirozenější aproximace standardní chyby $\sqrt{\text{E}[\tilde{\mathbf{x}}_k^T \tilde{\mathbf{x}}_k]}$ z konečného počtu realizací, která je úzce propojena se standardní odchylkou chyby odhadu. Pro skalární nestranný estimátor je to dokonce její nejlepší aproximace. Vzhledem k tomu, že standardní odchylka je velmi důležitý parametr pravděpodobnostní analýzy, tak je RMSE ve skalárním případě ideální pro analýzu výsledků simulace.

Nicméně, pohledem na výpočet RMSE zjistíme, že výsledná hodnota je silně ovlivněna velkými hodnotami $\|\tilde{\mathbf{x}}_k(i)\|$. Pro představu, kdyby 99 členů bylo okolo hodnoty 1 a jeden jediný kolem hodnoty 400, pak i výsledné RMSE by bylo kolem hodnoty 40. Proto je RMSE metrika pesimistická, neboť se příliš soustředí na to, jak špatný, byť i jediný, odhad je.

Popularita RMSE spočívá v příbuznosti s návrhovým kritériem MSE, které je pro své matematické vlastnosti (řešitelnost) nejpoužívanějším kritériem pro syntézu estimátorů. I když je tak RMSE často interpretováno, nejedná se o průměrnou vzdálenost v eukleidovském prostoru. Touto fyzikální interpretací se může pyšnit pouze metrika průměrné eukleidovské chyby uvedená v další části.

4.1.2 Průměrná eukleidovská chyba - AEE

Průměrná eukleidovská chyba (*Average Euclidean Error*) je definovaná jako

$$\text{AEE}(\hat{\mathbf{x}}_k) = \frac{1}{M} \sum_{i=1}^M \|\tilde{\mathbf{x}}_k(i)\|. \quad (4.2)$$

Jak i její název naznačuje, AEE vychází z konceptu eukleidovské vzdálenosti nebo také eukleidovské normy. Občas se o ní mluví (především ve skalárním případě) jako o střední absolutní chybě (*Mean Absolute Error*), nicméně tento název je pro vektory zavádějící a ještě k tomu může vést k domněnku, že jde o absolutní chybu ve smyslu opaku relativní chyby.

AEE je výběrový průměr skutečné střední hodnoty \bar{e} eukleidovské normy $\bar{e} = \text{E}[\|\tilde{\mathbf{x}}_k\|]$, a proto je také jeho nejlepší aproximací. Z tohoto důvodu má AEE spoustu cenných a žádoucích vlastností. Například AEE je nestranný odhad \bar{e} nezávisle na distribuci $\|\tilde{\mathbf{x}}_k\|$ a AEE je také nejlepší odhad \bar{e} ve smyslu nejmenších čtverců². Více v [21].

² $\sum_i (\|\tilde{\mathbf{x}}_k(i)\| - \hat{e})^2$ je minimalizováno pro $\hat{e} = \text{AEE}$

Podobně jako RMSE, stejně tak i AEE je pesimistická metrika. Obě tyto metriky tak příliš silně penalizují velké chyby a téměř ignorují malé chyby. Nicméně v příkladu z minulé sekce bude AEE přibližně 5 oproti hodnotě RMSE, která je zhruba 40, tedy AEE není tak pesimistické jako RMSE.

V mnoha aplikacích hodnocení kvality by tak AEE mohlo, a mělo být vhodnější metrikou [21], protože má jasnější fyzikální interpretaci a není tolik pesimistické. RMSE by ale bylo spolehlivější v pravděpodobnostní analýze (díky své vazbě na standardní odchylku).

4.1.3 Harmonický průměr chyb - HAE

V předchozích sekcích byly představeny pesimistické metriky RMSE a AEE. V některých situacích je lepší sledovat četnost malých chyb, které by vyvážily několik málo velkých chyb. Tehdy pomůže optimistická metrika. K tomuto účelu může být vhodný harmonický průměr chyb (harmonic average error)

$$\text{HAE}(\hat{\mathbf{x}}_k) = \left(\frac{1}{M} \sum_{i=1}^M \|\tilde{\mathbf{x}}_k(i)\|^{-1} \right)^{-1} = \frac{M}{\|\tilde{\mathbf{x}}_k(1)\|^{-1} + \dots + \|\tilde{\mathbf{x}}_k(M)\|^{-1}} \quad (4.3)$$

odvozený ze známého harmonického průměru. Výsledek HAE je nejvíce ovlivněn členy malých chyb a velké chyby jsou potlačovány. Toho lze s úspěchem využít v hodnocení kvality odhadu v systémech s velkým šumem, to jest v případech, kdy chyba odhadu silně kolísá pro jednotlivé iterace simulace Monte Carlo. Bohužel metrika HAE nemá jasnou fyzikální interpretaci, a tak se používá jen zřídka.

4.1.4 Geometrický průměr chyb - GAE

Ani jedna z předchozích metrik není vyvážená, proto může mít v některých situacích opodstatnění geometrický průměr chyb (*Geometric Average Error*) definovaný vztahem

$$\text{GAE}(\hat{\mathbf{x}}_k) = \left(\prod_{i=1}^M \|\tilde{\mathbf{x}}_k(i)\| \right)^{\frac{1}{M}} \quad (4.4)$$

nebo z výpočetních důvodů pomocí logaritmu

$$\ln [\text{GAE}(\hat{\mathbf{x}}_k)] = \frac{1}{M} \sum_{i=1}^M \ln \|\tilde{\mathbf{x}}_k(i)\|, \quad (4.5)$$

kde $\text{GAE}(\hat{\mathbf{x}}_k)$ je definováno jako 0, pokud je jedna nebo více eukleidovských norem rovna nule. Je zřejmé, že malá chyba může být vyvážena velkou chybou a naopak. Metrika je tak daleko méně ovlivněna extrémními hodnotami a spíše odpovídá typické hodnotě než AEE. Tato metrika je vhodná pro hodnocení poměrů norem nebo i stavů sestávajících převážně z indexů a koeficientů a proto bude používána v dalších sekcích.

4.1.5 Medián a modus chyby

Medián chyby je pro liché M prostřední člen pořádkové statistiky $\|\tilde{\mathbf{x}}_k(1)\|, \dots, \|\tilde{\mathbf{x}}_k(i)\|$. Pro sudé M pak aritmetický průměr prostředních dvou členů. Modus chyby je hodnota, kde má (vyhlazený) histogram získaný z $\|\tilde{\mathbf{x}}_k(1)\|, \dots, \|\tilde{\mathbf{x}}_k(M)\|$ své maximum.

Tyto dvě metriky nejsou ani optimistické ani pesimistické, neboť ani jednu z nich neovlivní velikost extrémních členů. V případě dostatečného vzorku dat tak reprezentují typickou chybu. Nicméně RMSE, AEE, GAE i HAE mají vlastnost, že přesnost jejich odhadů roste přímo úměrně s \sqrt{M} , a to bohužel neplatí pro medián ani modus. Medián je velmi důležitý v teorii chyby a ve finanční statistice. Naproti tomu modus je jediná metrika z uvedeného výčtu, která nám vždy dá nějakou reálnou chybu, která se vyskytuje v celém vzorku. Nejznámější aplikací modu je očekávaná délka dožití.

4.1.6 Iterační střední rozsah chyby - IMRE

V článku [43] jsou definovány tzv. robustní metriky centrální tendence. Jde o metriky, které jsou v určitých ohledech robustní, a zároveň vhodným způsobem vystihují hodnotu, kolem které jsou chyby odhadu distribuovány, tedy centrální tendenci. Autor si tamtéž zároveň oponuje, že jeho požadavky nesplňuje žádná současná metrika beze zbytku (kritéria a důkazy lze nalézt v [43]). Proto navrhuje novou metriku: iterační střední rozsah chyby (*Iterative Mid-Range Error*, IMRE) vycházející ze středního rozsahu chyby $\text{Mid-Range}\{S_M\}$. Střední rozsah je definován jako aritmetický průměr nejvyšší a nejnižší hodnoty množiny $S_M \triangleq \{\|\tilde{\mathbf{x}}_k(1)\|, \dots, \|\tilde{\mathbf{x}}_k(M)\|\}$, tj.

$$\text{Mid-Range}\{S_M\} = \frac{\max S_M - \min S_M}{2}. \quad (4.6)$$

Protože je střední rozsah chyby ovlivněn pouze dvěma extrémními členy množiny S_M , bylo představeno vylepšení připomínající svým algoritmem metodu půlení intervalů pro aproximativní řešení rovnic. Toto vylepšení, IMRE, je definované v jednom iteračním kroku jako nahrazení minimálního a maximálního členu jejich středním rozsahem, tj.

$$S_{M-k-1}^{(k+1)} = S_{M-k}^{(k)} + \{\text{Mid-Range}\{S_{M-k}^{(k)}\}\} - \{\min\{S_{M-k}^{(k)}\}\} - \{\max\{S_{M-k}^{(k)}\}\}, \quad (4.7)$$

pro k značící číslo iterace jdoucí od 0 do $M - 2$. Výsledkem je poslední zbývající člen množiny.

Pro lepší představu bude uveden výpočet IMRE z množiny chyb $S_4^{(0)} = \{1, 2, 3, 7\}$. Střední rozsah S_4^0 je $\text{Mid-Range}\{S_4^{(0)}\} = \frac{1+7}{2} = 4$. Nová množina bude $S_3^{(1)} = \{2, 3, 4\}$. K další iteraci je třeba znát střední rozsah S_3^1 , který je $\text{Mid-Range}\{S_3^{(1)}\} = \frac{2+4}{2} = 3$. Nová množina bude $S_2^{(2)} = \{3, 3\}$, nyní je již zřejmé, že po třetí iteraci nakonec zůstane poslední člen $S_1^{(3)} = \{3\}$, který je výslednou hodnotou IMRE.

4.1.7 Zobecnění metrik a poznámky

Zobecněním některých výše uvedených metrik je tzv. spektrum chyby (Error spectrum), které bylo prvně představeno v [23]. Je definováno pro absolutní chybu odhadu $e = \|\tilde{\mathbf{x}}_k\|$ nebo relativní chybu odhadu $e = \frac{\|\tilde{\mathbf{x}}_k\|}{\|\mathbf{x}\|}$ vztahem

$$S(r) = [E(e^r)]^{1/r}. \quad (4.8)$$

Spektrum chyby má mnoho zajímavých vlastností. Například $S(2) = \text{RMSE}$, $S(1) = \text{AEE}$, $S(0) = \text{GAE}$, $S(-1) = \text{HAE}$, $S(\infty) = \max e$ a $S(-\infty) = \min e$. $S(r)$ je také rostoucí a pro $r < 0$ je spektrum optimistické, pro $r = 0$ vyvážené a pro $r > 0$ pesimistické. Z výše uvedeného platí nerovnost

$$\text{HAE} \leq \text{GAE} \leq \text{AAE} \leq \text{RMSE}. \quad (4.9)$$

Článek [43] dokonce propojuje zdánlivě velmi odlišné metriky AEE, medián a modus. Definováním tzv. zobecněné vážené průměrné eukleidovské chyby (*Generalized Weighted AEE*, GW-AEE) jako

$$\text{GW-AEE}(\hat{\mathbf{x}}_k) = \sum_{i=1}^M \lambda_i \|\tilde{\mathbf{x}}_k(i)\|, \quad (4.10)$$

kde $\lambda_i \geq 0$ je váhová konstanta s vlastností $\sum_{i=1}^M \lambda_i = 1$. Pak jsou výše uvedené metriky vyjádřeny v rámci GW-AEE tímto:

- AEE

$$\lambda_i = \frac{1}{M},$$

- medián

$$\lambda_i = \begin{cases} 1 & \text{pokud } \|\tilde{\mathbf{x}}_k(i)\| \text{ je prostřední člen při lichém } M, \\ 0.5 & \text{pokud } \|\tilde{\mathbf{x}}_k(i)\| \text{ jsou prostřední dva členy při sudém } M, \\ 0 & \text{jinak.} \end{cases}$$

- modus

$$\lambda_i = \begin{cases} 1 & \text{pokud } \|\tilde{\mathbf{x}}_k(i)\| \text{ je vrcholem (vyhlazeného) histogramu,} \\ 0 & \text{jinak.} \end{cases}$$

Přesto jsou spektrum chyby a GW-AEE dvě rozdílné metriky agregující různé metriky vyjma AEE, které je agregováno v obou zobecněních.

Při aplikaci jakékoliv z představených metrik je třeba mít na paměti, že mohou nastat problémy s různorodostí stavových proměnných podrobně vysvětlených v sekci 4.4. Proto je vhodné před výpočtem zvážit tamtéž nastíněné řešení.

4.2 Metriky relativní chyby

Veškeré metriky z předchozí části jsou absolutní, protože nejsou vztaženy k žádné referenční hodnotě. Tyto metriky jsou citlivé na podmínky experimentu, například změna počáteční podmínky může vyvolat velké změny hodnot těchto metrik. To nevadí, pokud je cílem hodnotit celý systém včetně estimátoru jako celek, pokud ale jde o hodnocení estimačních algoritmů, pak je lepší použít metriky relativní chyby, aby byly výsledky vůbec porovnatelné.

Metriky relativní chyby jsou vždy vztaženy k nějaké referenční hodnotě. Takovýchto referenčních hodnot je mnoho, ale prakticky se používají tyto:

- norma vektoru skutečného stavu $\|\mathbf{x}_k\|$,
- chyba měření,
- chyba prediktivního odhadu značená jako $\|\mathbf{x}_k - \hat{\mathbf{x}}_k^*\|$, kde $\hat{\mathbf{x}}_k^*$ označuje prediktivní odhad $\hat{\mathbf{x}}_{k|k-1}$.

Když máme zvolenou referenční hodnotu, můžeme si vybrat, jakým způsobem ji použijeme k výpočtu relativní chyby. Nejpřirozenější je buď

$$\frac{\text{průměrná chyba odhadu}}{\text{průměrná referenční hodnota}} \quad \text{nebo} \quad \text{průměr} \left(\frac{\text{chyba odhadu}}{\text{referenční hodnota}} \right),$$

přičemž druhá možnost má obecně jasnější interpretaci a menší varianci.

4.2.1 Relativní verze vybraných metrik absolutní chyby

Relativní metriky zkonstruované z absolutních metrik z předchozí sekce jsou podle výše uvedeného rozděleny do dvou sloupců na

$$\frac{\text{metrika chyby odhadu}}{\text{metrika stavu}} \quad \text{nebo} \quad \text{metrika} \left(\frac{\text{chyba odhadu}}{\text{norma stavu}} \right)$$

a uvedeny v následujícím výčtu:

- relativní odmocnina ze střední kvadratické chyby (*Root Mean Squared Relative Error*, RMSRE)

$$\frac{\sqrt{\frac{1}{M} \sum_{i=1}^M \|\tilde{\mathbf{x}}_k(i)\|^2}}{\sqrt{\frac{1}{M} \sum_{i=1}^M \|\mathbf{x}_k(i)\|^2}} \quad \text{nebo} \quad \sqrt{\frac{\frac{1}{M} \sum_{i=1}^M \|\tilde{\mathbf{x}}_k(i)\|^2}{\frac{1}{M} \sum_{i=1}^M \|\mathbf{x}_k(i)\|^2}}$$

- průměrná relativní eukleidovská chyba (*Average Relative Euclidean Error*, ARE)

$$\frac{\frac{1}{M} \sum_{i=1}^M \|\tilde{\mathbf{x}}_k(i)\|}{\frac{1}{M} \sum_{i=1}^M \|\mathbf{x}_k(i)\|} \quad \text{nebo} \quad \frac{1}{M} \sum_{i=1}^M \frac{\|\tilde{\mathbf{x}}_k(i)\|}{\|\mathbf{x}_k(i)\|},$$

- průměrná geometrická relativní chyba (*Geometric Average Relative Error*, GRE)

$$\frac{1}{M} \sum_{i=1}^M \ln \frac{\|\tilde{\mathbf{x}}_k(i)\|}{\|\mathbf{x}_k(i)\|}$$

- průměrná harmonická relativní chyba (*Harmonic Average Relative Error*, HRE)

$$\frac{\left(\frac{1}{M} \sum_{i=1}^M \|\tilde{\mathbf{x}}_k(i)\|^{-1} \right)^{-1}}{\left(\frac{1}{M} \sum_{i=1}^M \|\mathbf{x}_k(i)\|^{-1} \right)^{-1}} \quad \text{nebo} \quad \left(\frac{1}{M} \sum_{i=1}^M \frac{\|\tilde{\mathbf{x}}_k(i)\|^{-1}}{\|\mathbf{x}_k(i)\|^{-1}} \right)^{-1}$$

- medián relativní chyby (*Median Relative Error*, MRE)

$$\frac{\text{medián } z \{ \|\tilde{\mathbf{x}}_k(1)\|, \dots, \|\tilde{\mathbf{x}}_k(M)\| \}}{\text{medián } z \{ \|\mathbf{x}_k(1)\|, \dots, \|\mathbf{x}_k(M)\| \}} \text{ nebo medián } z \left\{ \frac{\|\tilde{\mathbf{x}}_k(1)\|}{\|\mathbf{x}_k(1)\|}, \dots, \frac{\|\tilde{\mathbf{x}}_k(M)\|}{\|\mathbf{x}_k(M)\|} \right\},$$

- modus relativní chyby (*Relative Error Mode*, REM)

$$\frac{\text{modus } z \{ \|\tilde{\mathbf{x}}_k(1)\|, \dots, \|\tilde{\mathbf{x}}_k(M)\| \}}{\text{modus } z \{ \|\mathbf{x}_k(1)\|, \dots, \|\mathbf{x}_k(M)\| \}} \text{ nebo modus } z \left\{ \frac{\|\tilde{\mathbf{x}}_k(1)\|}{\|\mathbf{x}_k(1)\|}, \dots, \frac{\|\tilde{\mathbf{x}}_k(M)\|}{\|\mathbf{x}_k(M)\|} \right\}.$$

Tyto metriky mají vesměs stejné výhody a nevýhody jako jejich absolutní verze. Nej-slibněji ze všech se jeví GRE, protože její základ, metrika GAE je nejlepší pro zpracování poměrů, což je přesně případ relativních metrik.

Nyní bude uvedeno několik relativních metrik představených v [20].

4.2.2 Kvocient chyby bayesovského odhadu - BEEQ

Kvocient chyby bayesovského odhadu (*Bayesian Estimation Error Quotient*, BEEQ) je metrika relativní chyby odhadu, která kvantifikuje zlepšení filtračního odhadu $\hat{\mathbf{x}}_k$ oproti prediktivnímu odhadu $\hat{\mathbf{x}}_k^*$, definovaná jako

$$r^*(\hat{\mathbf{x}}_k) = \frac{\text{AEE}(\hat{\mathbf{x}}_k)}{\text{AEE}(\hat{\mathbf{x}}_k^*)} = \frac{\sum_{i=1}^M \|\mathbf{x}_k(i) - \hat{\mathbf{x}}_k(i)\|}{\sum_{i=1}^M \|\mathbf{x}_k(i) - \hat{\mathbf{x}}_k^*(i)\|}, \quad (4.11)$$

kde $\hat{\mathbf{x}}_k(i)$ a $\hat{\mathbf{x}}_k^*(i)$ jsou filtrační, resp. prediktivní odhad stavu \mathbf{x} v i -té iteraci simulace Monte Carlo. Podobně jako u ostatních relativních metrik lze k jejich definici přistoupit dvěma způsoby. Buď poměrem dvou absolutních metrik jako v (4.11) nebo průměrem z poměru dvou norem

$$r_k(i) = \frac{\|\mathbf{x}_k(i) - \hat{\mathbf{x}}_k(i)\|}{\|\mathbf{x}_k(i) - \hat{\mathbf{x}}_k^*(i)\|}. \quad (4.12)$$

Pak je ale vhodnější upřednostnit před aritmetickým průměrem v (4.11) průměr geometrický (pro jeho vyváženost, tj. schopnost rozumně potlačovat výjimečně velké hodnoty) a zároveň definovat BEEQ z výpočetních důvodů pomocí logaritmu jako

$$\ln[r(\hat{\mathbf{x}}_k)] = \frac{1}{M} \sum_{i=1}^M \ln r_k(i). \quad (4.13)$$

BEEQ kvantifikuje přínos nově přichozích měření pro bayesovský rekursivní odhad. Z logiky věci plyne, že dobrý estimátor bude mít BEEQ významně nižší než 1, protože chyba filtračního odhadu $\|\mathbf{x}_k - \hat{\mathbf{x}}_k\|$ by vždy měla být nižší než chyba prediktivního odhadu $\|\mathbf{x}_k - \hat{\mathbf{x}}_k^*\|$. Pokud se BEEQ bude držet na zhruba konstantní hodnotě, lze tvrdit, že přínos nových měření není již tak vysoký, protože model systému je již dobře nakalibrován a chyba prediktivního odhadu je minimální. Odhady se tedy blíží k mezi poznatelnosti. Naopak významný pokles hodnoty BEEQ indikuje nárůst významu měření s ohledem na velké chyby prediktivního odhadu.

4.2.3 Chyba odhadu vztažená k chybě měření - EMER

Další metrika vychází z úvahy, že přesnost odhadu kriticky závisí na přesnosti měření. Ke kvantifikaci přínosu měření pro rekurzivní odhad byla navržena metrika poměr chyby odhadu a měření (*Estimate-Measurement Error Ratio*, EMER) definovaná jako

$$\rho^*(\hat{\mathbf{x}}_k) = \frac{\text{AEE}^z(\mathbf{h}_k(\hat{\mathbf{x}}_k))}{\text{AEE}^z(\mathbf{z}_k)} = \frac{\sum_{i=1}^M \|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{h}_k(\hat{\mathbf{x}}_k(i))\|}{\sum_{i=1}^M \|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{z}_k(i)\|}. \quad (4.14)$$

Vztah (4.14) je poměr dvou AEE metrik v prostoru měření. Podobně jako v případě BEEQ může být někdy vhodnější využít geometrických poměrů samostatných EMER, tj.

$$\ln[\rho(\hat{\mathbf{x}}_k)] = \frac{1}{M} \sum_{i=1}^M \ln \rho_k(i), \quad (4.15)$$

kde

$$\rho_k(i) = \frac{\|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{h}_k(\hat{\mathbf{x}}_k(i))\|}{\|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{z}_k(i)\|}. \quad (4.16)$$

Nicméně zde porovnáváme kvalitu odhadu stavu v prostoru měření, nikoliv ve stavovém prostoru. V situaci, kdy bude existovat inverze \mathbf{h}_k^{-1} , by měla být preferována verze EMER ve stavovém prostoru, tj.

$$\rho^*(\hat{\mathbf{x}}_k) = \frac{\sum_{i=1}^M \|\mathbf{x}_k(i) - \hat{\mathbf{x}}_k(i)\|}{\sum_{i=1}^M \|\mathbf{x}_k(i) - \mathbf{h}_k^{-1}(\mathbf{z}_k(i))\|}, \quad (4.17)$$

resp.

$$\ln[\rho(\hat{\mathbf{x}}_k)] = \frac{1}{M} \sum_{i=1}^M \ln \rho_k^x(i), \quad (4.18)$$

kde

$$\rho_k^x(i) = \frac{\|\mathbf{x}_k(i) - \hat{\mathbf{x}}_k(i)\|}{\|\mathbf{x}_k(i) - \mathbf{h}_k^{-1}(\mathbf{z}_k(i))\|}. \quad (4.19)$$

Podobně jako v případě BEEQ i u EMER můžeme pro dobrý estimátor očekávat $\text{EMER} < 1$. Čím nižší EMER, tím lepší redukce chyby. Jinak by triviální volba $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^z$ splňující rovnost $\mathbf{h}_k(\hat{\mathbf{x}}_k^z) = \mathbf{z}_k$ zaručovala $\text{EMER} = 1$. Další informace k této metrice včetně verze EMER pro porovnání estimátorů pracujících s různou dimenzí měření lze nalézt v [21].

4.2.4 Redukce chyby bayesovského odhadu - BERF

Metrika faktor redukce bayesovské chyby (*Bayesian Error Reduction Factor*, BERF) je kompromisem mezi BEEQ a EMER kvantifikujícím celkové zlepšení odhadu v jednom filtračním a prediktivním kroku zároveň. Je definována jako

$$\eta^*(\hat{\mathbf{x}}_k) = \frac{\text{BEEQ}^z(\hat{\mathbf{x}}_k) + \beta_k \cdot \text{EMER}(\hat{\mathbf{x}}_k)}{1 + \beta_k}, \quad (4.20)$$

kde $\text{BEEQ}^z(\hat{\mathbf{x}}_k)$ je BEEQ v prostoru měření definované

$$\text{BEEQ}^z(\hat{\mathbf{x}}_k) = \frac{\text{AEE}^z(\mathbf{h}_k(\hat{\mathbf{x}}_k))}{\text{AEE}^z(\mathbf{h}_k(\hat{\mathbf{x}}_k^*))} = \frac{\sum_{i=1}^M \|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{h}_k(\hat{\mathbf{x}}_k(i))\|}{\sum_{i=1}^M \|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{h}_k(\hat{\mathbf{x}}_k^*(i))\|}, \quad (4.21)$$

EMER($\hat{\mathbf{x}}_k$) je definované vztahem (4.14) a koeficient β_k je definován

$$\beta_k = \frac{\text{AEE}^z(\mathbf{h}_k(\hat{\mathbf{x}}_k^*))}{\text{AEE}^z(\mathbf{z}_k)} = \frac{\sum_{i=1}^M \|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{h}_k(\hat{\mathbf{x}}_k^*(i))\|}{\sum_{i=1}^M \|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{z}_k(i)\|}. \quad (4.22)$$

Nebo opět lépe pomocí geometrického průměru a logaritmu individuálních BERF:

$$\ln[\eta(\hat{\mathbf{x}}_k)] = \frac{1}{M} \sum_{i=1}^M \ln \eta_k(i), \quad (4.23)$$

$$\eta_k(i) = \frac{r_k^z(i) + \beta_k(i)\rho_k(i)}{1 + \beta_k(i)}, \quad (4.24)$$

kde jednotlivé individuální proměnné jsou

$$r_k^z(i) = \frac{\|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{h}_k(\hat{\mathbf{x}}_k(i))\|}{\|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{h}_k(\hat{\mathbf{x}}_k^*(i))\|}, \quad \beta_k(i) = \frac{\|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{h}_k(\hat{\mathbf{x}}_k(i)^*)\|}{\|\mathbf{h}_k(\mathbf{x}_k(i)) - \mathbf{z}_k(i)\|} \quad (4.25)$$

a $\rho_k(i)$ je dáno vztahem (4.16).

Jak již bylo řečeno výše BERF, je kompromisem mezi BEEQ a EMER, z jehož definice vyplývá, že hodnota BERF bude vždy ležet mezi hodnotou EMER a BEEQ. Pokud v systému odhadu bude převažovat chyba filtračního odhadu, bude BERF tíhnout k hodnotě EMER. Pokud naopak bude převažovat chyba predikce, bude BERF tíhnout k hodnotě BEEQ. Další informace lze opět nalézt v [21].

4.2.5 Poznámky

Pro všechny relativní metriky platí totéž co pro absolutní metriky uvedené v sekci 4.4, a je vhodné se těmito doporučeními řídit. Novější metriky BEEQ, EMER a BERF nejsou pravděpodobně moc používané, nicméně dle jejich autora se může jednat o zajímavé nástroje pro hodnocení kvality odhadu různých estimátorů. Požadavky pro dobrý estimátor jsou jasné, tyto tři metriky by měly být menší než 1. Čím nižší, tím lépe.

4.3 Míry četnosti

Doposud představené metriky určitým způsobem měří velikost chyby odhadu. V některých aplikacích je ale vhodnější měřit, zda a jak často je odhad přijatelný v rámci nějaké tolerance. Typickým příkladem je navádění rakety na cíl. Pokud raketa mine svůj cíl, je již v podstatě jedno o kolik. Proto je jistě zajímavější sledovat, kolikrát ze 100 pokusů raketa cíl trefí. Metrikám tohoto typu říkáme míry četnosti (nějakých událostí) a budou představeny v následujících odstavcích.

Stěžejním prvkem metrik toho typu je definování oblasti (ne)úspěšnosti. Takováto oblast je jistě závislá na dané aplikaci, ale většinou se jako oblast používá hyperkoule

se středem v bodě \mathbf{x}_k a nějakým poloměrem. X. Rong Li [21] si všímá, že by tato oblast měla být funkcí měření (nelze očekávat dobrou úspěšnost při nekvalitních vstupních datech). Proto v případě, že není oblast (ne)úspěšnosti jasná ze zadání problému, doporučuje používat

$$R = \{\hat{\mathbf{x}}_k(i) : \rho_k(i)(\hat{\mathbf{x}}_k(i)) \leq \delta\}, \quad (4.26)$$

kde $\rho_k(i)$ je individuální EMER definované vztahem (4.16). Tato oblast je definována na základě relativní metriky, proto se hodí spíše pro hodnocení kvality odhadu různých estimátorů.

4.3.1 Míra úspěšnosti

Oblast úspěšnosti (success region) je definován jako

$$R_s = \{\hat{\mathbf{x}}_k(i) : \rho_k(i)(\hat{\mathbf{x}}_k(i)) \leq \delta_s\}, \quad (4.27)$$

pro malé δ_s (obvykle $\delta_s < 1$).

Míra úspěšnosti estimátoru je procento odhadů, jež padnou dovnitř regionu úspěšnosti. Její skutečná hodnota se počítá aproximací skutečné pravděpodobnosti $P\{\hat{\mathbf{x}}_k \in R_s\}$ konečným počtem realizací.

Někdy je zajímavé hledat naopak takové δ_s , aby uvnitř regionu úspěšnosti leželo požadované procento odhadů, pak mluvíme o percentilu koncentrace.

4.3.2 Míra neúspěšnosti

Oblast akceptovatelnosti (feasible region) je definována jako

$$R_f = \{\hat{\mathbf{x}}_k(i) : \rho_k(i)(\hat{\mathbf{x}}_k(i)) \leq \delta_f\}, \quad (4.28)$$

pro velké δ_f (obvykle $\delta_f > 1$).

Míra neúspěšnosti estimátoru je procento odhadů, jež padnou mimo region úspěšnosti. Její skutečná hodnota se počítá aproximací skutečné pravděpodobnosti $P\{\hat{\mathbf{x}}_k \notin R_f\}$ konečným počtem realizací.

4.3.3 Poznámky

Míra úspěšnosti je optimistickou metrikou, protože hodnotí jak moc je estimátor dobrý (špatné výsledky zcela ignoruje), naproti tomu míra neúspěšnosti je z opačných důvodů pesimistickou metrikou. Tyto dvě metriky nemusí být vzájemně komplementární z důvodu obecně odlišných hodnot δ_s a δ_f . Míra úspěšnosti by měla preferovat estimátory založené na maximální aposteriorní pravděpodobnosti (MAP) a maximální věrohodnosti (ML). Naproti tomu míra neúspěšnosti by měla vycházet lépe pro estimátory minimalizující střední kvadratickou chybu (MMSE).

4.4 Problém fyzikálně různorodých stavů a stavů řádově odlišných hodnot

Základním úkolem všech metrik je promítnout rozdíl mezi skutečným stavem a odhadem stavu do jediné hodnoty. Přitom by mělo platit, že čím větší je tento rozdíl, tím vyšší by měla být výsledná hodnota metriky. Problém získání jediného čísla postihujícího celkovou chybu $\tilde{\mathbf{x}}_k$ se obvykle řeší normováním chyby, tj. výpočtem nějaké normy této chyby, z jednotlivých iterací $\tilde{\mathbf{x}}_k(i)$ simulace MC a jejich následným průměrováním. Při normování ale dochází ke ztrátě mnoha informací.

Například v úloze navigace se běžně ve stavu vyskytují stavové proměnné reprezentující polohu, rychlost a někdy dokonce i zrychlení. Na otázku, co vznikne normováním navzájem nesouvisejících hodnot, lze těžko odpovědět. Bude to pouze nějaké číslo bez zjevné fyzikální interpretace.

Dalším problémem může být normování řádově odlišných hodnot. Například za předpokladu skutečného stavu $\mathbf{x} = [1, 1, 1, 1000]^T$ a odhadu $\hat{\mathbf{x}} = [2, 1, 1, 1010]^T$ nelze očekávat, že by norma chyby stavu $\tilde{\mathbf{x}}_k$ odhalila o 100 % nadhodnocený odhad první složky stavu.

Praktickým řešením problémů z předchozích odstavců je rozdělení vektoru stavu na fyzikálně, či alespoň řádově porovnatelné složky. Právě v již zmíněné úloze navigace se tento problém řeší elegantně rozdělením stavu na složky polohy a rychlosti. Hodnoty metrik jsou pak vypočteny pro tyto samostatné složky stavu a výsledek je jednoduše fyzikálně interpretovatelný.

Kapitola 5

Metriky kovarianční a MSE matice

Tato kapitola prezentuje metriky¹, které využívají druhého momentu náhodných veličin. To je velký rozdíl oproti metrikám z minulé kapitoly, které využívají pouze první moment - chybu odhadu $\tilde{\mathbf{x}}_k$. Využitím druhého momentu je doufáno v extrahování více informací o kvalitě odhadu. Tímto druhým momentem je buď matice středních kvadratických chyb (MSE matice) nebo kovarianční matice chyby odhadu $\mathbf{P}_{k|k}$, která bude pro zjednodušení zápisu označována \mathbf{P}_k . Kovarianční matice \mathbf{P}_k vyjadřuje vlastní hodnocení chyby odhadu estimátorem. Matice střední kvadratické chyby Σ_k vyjadřuje střední kvadratické chyby odchylky odhadu, tj.

$$\Sigma_k = E[\tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^T] = E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T]. \quad (5.1)$$

Protože se jedná o střední hodnotu, která je málokdy známá, stanovuje se většinou její odhad na základě konečného počtu realizací (M) Monte Carlo simulací

$$\Sigma_k \approx \frac{1}{M} \sum_{i=1}^M \tilde{\mathbf{x}}_k(i) \tilde{\mathbf{x}}_k^T(i). \quad (5.2)$$

5.1 Hodnocení samostatné matice

V některých aplikacích (např. navigace či sledování pohybujícího se objektu) je zvykem vyhodnotit kromě absolutní chyby odhadu také samostatnou kovarianční matici \mathbf{P}_k a spolu s odhadem stavu $\hat{\mathbf{x}}_k$ z toho činit společné závěry.

Kovarianční matice \mathbf{P}_k obsahuje $\frac{n_x+1}{2}$ -krát více hodnot než stav \mathbf{x}_k nebo odhad stavu $\hat{\mathbf{x}}_k$. Důsledkem je, že samostatné skalární ohodnocení této matice je větší kompromis než ohodnocení vektoru. Povahu jakékoliv matice nejlépe vystihují její vlastní čísla, tedy kořeny charakteristického polynomu matice. Z tohoto důvodu patří mezi nejpoužívanější maticová hodnocení téměř výlučně ta hodnocení, která jsou založena na analýze vlastních čísel.

Mezi nejvýznamnější hodnocení kovarianční matice \mathbf{P}_k patří:

¹Opět jako v minulé kapitole je termín metrika chápán volně jako hodnota oceňující kvalitu odhadu, nikoliv v striktně matematickém pojetí. V anglicky psané odborné literatuře se termín „metric“ běžně pro tento účel používá.

- stopa

$$\operatorname{tr} \mathbf{P}_k = \sum_{j=1}^{n_x} \lambda_j, \quad (5.3)$$

- determinant

$$\det \mathbf{P}_k = \prod_{j=1}^{n_x} \lambda_j, \quad (5.4)$$

- minimální nebo maximální vlastní číslo

$$\min\{\lambda_1, \dots, \lambda_{n_x}\} \text{ nebo } \max\{\lambda_1, \dots, \lambda_{n_x}\}, \quad (5.5)$$

- Frobeniova maticová norma

$$\|\mathbf{P}_k\|_F = \sqrt{\sum_{i=1}^{n_x} \sum_{j=1}^{n_x} |p_k^{ij}|^2} = \sqrt{\operatorname{tr}(\mathbf{P}_k^T \mathbf{P}_k)} = \sqrt{\sum_{j=1}^{n_x} \lambda_j^2}, \quad (5.6)$$

kde λ_j značí j -té vlastní číslo matice \mathbf{P}_k a p_k^{ij} je prvek matice \mathbf{P}_k , který se nachází na i -tém řádku a j -tém sloupci.

Přestože je vlastních čísel stejně jako je složek stavu, nelze mezi nimi najít jasné propojení a přiřadit každé jedno vlastní číslo některé ze složek stavu. To znesnadňuje interpretaci těchto hodnocení a tato hodnocení doplácí na problémy fyzikálně či řádově různorodých stavů popsaných v sekci 4.4. Řešení je obdobné tomu nastíněnému ve zmíněné sekci, tj. rozdělit kovarianční matici \mathbf{P}_k na části odpovídající si fyzikálně či alespoň řádově a hodnotit až tyto submatice.

Vraťme se k typickému nasazení těchto hodnocení z úvodu této části, k úloze sledování pohybujícího se objektu v n_x -rozměrném prostoru. Při hodnocení kvality odhadu kvality je kovarianční matice \mathbf{P}_k nejprve rozdělena na submatice $n \times n$ odpovídající poloze, rychlosti a případně i zrychlení sledovaného objektu. Na tyto submatice jsou následně aplikována výše představená hodnocení matic a z výsledných hodnot jsou vyvozeny závěry o rozprostření distribuce aposteriori odhadu. Za předpokladu gaussovského rozdělení reprezentuje kovarianční submatice polohy elipsoid obklopující odhad polohy $\hat{\mathbf{x}}_k$, jehož pravděpodobnost překročí nějakou prahovou hodnotu. Délka os tohoto elipsoidu je proporcionální vlastním číslům \mathbf{P}_k . Fyzikální a geometrická interpretace hodnocení matic je následující [42]:

- stopa matice vyjadřuje celkovou očekávanou střední kvadratickou chybu, zároveň je to součet délky os elipsoidu,
- determinant má spojitost s objemem elipsoidu,
- minimální a maximální vlastní čísla definují minimální nebo maximální chybu v jedné ose elipsoidu,
- Frobeniova norma je ve své podstatě rozšířením eukleidovské normy na matice, tedy součet kvadrátů délky os elipsoidu.

5.2 Věrohodnost odhadu

Většina estimátorů využívá k rekurzivnímu výpočtu odhadu stavu $\hat{\mathbf{x}}_k$ kovarianční matice z minulého časového okamžiku \mathbf{P}_{k-1} . Je nasnadě, že čím vyšší kvality bude tato kovarianční matice vykazovat, tím bude odhad přesnější. Vyšší kvalita je většinou chápána ve smyslu přiblížení se kovarianční matice (reprezentující vlastní hodnocení chyby estimátoru) k matici skutečné střední kvadratické chyby Σ_k . Jde tedy o zkoumání věrohodnosti kovarianční matice.

V práci [24] je věrohodnost definována následovně:

Estimátor je věrohodný na hladině α ($0 \leq \alpha \leq 1$), pokud je rozdíl mezi kovarianční maticí \mathbf{P}_k a maticí střední kvadratické chyby Σ_k statisticky nevýznamný na hladině α . O maximální hladině α , na které se věrohodnost estimátoru zamítá, se mluví jako o hladině nevěrohodnosti estimátoru. Estimátor je optimistický (příliš sebevědomý) na hladině α , jestliže jeho sebehodnocení je statisticky menší než skutečná matice MSE (nebo výchylka) na hladině α . Estimátor je pesimistický (příliš nesmělý) v opačném případě.

Ve starších pracích [4, 5] se používá termín konzistence (*finite - sample consistency*) namísto věrohodnosti. To může být dle [24] zavádějící, proto je pojem věrohodnost preferován i v této práci.

V této práci je zkoumána věrohodnost za předpokladu nestrannosti estimátoru.

5.2.1 Jednoduchá hodnocení věrohodnosti

Za zmíněného předpokladu nestrannosti estimátoru je problém vztažen pouze na porovnání MSE matice Σ_k a kovarianční matice \mathbf{P}_k . Pod pojmem jednoduchá hodnocení si lze představit jakékoliv operace zohledňující rozdíl mezi právě těmito dvěma maticemi jako jedinou hodnotu. Protože pro porovnání dvou matic neexistuje žádný ucelený postup, lze vymyslet nepřeberné množství metrik. Proto bude uveden spíše výčet možností, které se naskýtají [24, 25].

Jednou z nejzajímavějších možností je vyšetřit věrohodnost pomocí stopy matice, protože stopa $\text{tr}(\Sigma_k)$ je součet kvadrátů chyb odhadu a stopa $\text{tr}(\mathbf{P}_k)$ je odhad součtu kvadrátů chyb dodaný estimátorem. Nabízí se tak například absolutní a relativní metriky typu MSER (*Mean Square Error Ratio*):

$$\text{MSER}_k^1 = \frac{\text{tr} \mathbf{P}_k}{\text{tr} \Sigma_k}, \quad \text{MSER}_k^2 = \log \frac{\text{tr} \Sigma_k}{\text{tr} \mathbf{P}_k}, \quad (5.7)$$

$$\text{MSER}_k^3 = \frac{\text{tr} \Sigma_k - \text{tr} \mathbf{P}_k}{\text{tr} \Sigma_k} = \frac{\text{tr}(\Sigma_k - \mathbf{P}_k)}{\text{tr} \Sigma_k}, \quad (5.8)$$

$$\text{MSER}_k^4 = \frac{\text{tr} \Sigma_k - \text{tr} \mathbf{P}_k}{\text{tr} \Sigma_k + \text{tr} \mathbf{P}_k} = \frac{\text{tr}(\Sigma_k - \mathbf{P}_k)}{\text{tr}(\Sigma_k + \mathbf{P}_k)}. \quad (5.9)$$

Přestože lze tato hodnocení poměrně dobře interpretovat, je zřejmé, že závisí pouze na diagonálních prvcích. Proto může být někdy vhodnější použít hodnocení založená na determinantu - GEVR (*Generalized Error Variance Ratio*). Tato hodnocení mají díky

použití determinantu vztah k objemu rovnoběžnostěnu a lze uvést například:

$$\text{GEVR}_k^1 = \frac{\det \mathbf{P}_k}{\det \boldsymbol{\Sigma}_k}, \quad \text{GEVR}_k^2 = \log \frac{\det \boldsymbol{\Sigma}_k}{\det \mathbf{P}_k}, \quad (5.10)$$

$$\text{GEVR}_k^3 = \frac{\det \boldsymbol{\Sigma}_k - \det \mathbf{P}_k}{\det \boldsymbol{\Sigma}_k}, \quad \text{GEVR}_k^4 = \frac{\det \boldsymbol{\Sigma}_k - \det \mathbf{P}_k}{\det \boldsymbol{\Sigma}_k + \det \mathbf{P}_k}, \quad (5.11)$$

Pokud je žádoucí porovnat skutečně všechny prvky, nabízí se spíše hodnocení založená na Frobeniově (nebo jiné) maticové normě - MNR (*Matrix Norm Ratio*):

$$\text{MNR}_k^1 = \frac{\|\mathbf{P}_k\|}{\|\boldsymbol{\Sigma}_k\|}, \quad \text{MNR}_k^2 = \log \frac{\|\boldsymbol{\Sigma}_k\|}{\|\mathbf{P}_k\|} \quad (5.12)$$

$$\text{MNR}_k^3 = \frac{\|\boldsymbol{\Sigma}_k\| - \|\mathbf{P}_k\|}{\|\boldsymbol{\Sigma}_k + \mathbf{P}_k\|}, \quad \text{MNR}_k^4 = \frac{\|\boldsymbol{\Sigma}_k - \mathbf{P}_k\|}{\|\boldsymbol{\Sigma}_k + \mathbf{P}_k\|}. \quad (5.13)$$

Na tomto místě je opět vhodné připomenout, že výše představená hodnocení jsou pouze různé návrhy vyhodnocení rozdílu mezi kovarianční maticí \mathbf{P}_k a maticí $\boldsymbol{\Sigma}_k$. Při aplikaci jednoduchého hodnocení věrohodnosti je tak vždy nutné hledat optimální řešení plynoucí ze zadaného problému a případně nějakým způsobem zprůměrovat kovarianční matice z různých iterací Monte Carlo.

5.2.2 Normalizovaný kvadrát chyby odhadu - NEES

Vůbec první metoda k vyhodnocení věrohodnosti odhadu poskytovaného estimátorem byla založena na normalizovaném kvadrátu chyby odhadu (*Normalized Estimation Error Squared*, NEES) představeném v [4] a definovaným jako

$$\epsilon_k(i) = \tilde{\mathbf{x}}_k^T(i) \mathbf{P}_k^{-1}(i) \tilde{\mathbf{x}}_k(i) = \|\tilde{\mathbf{x}}_k(i)\|_{\mathbf{P}_k^{-1}(i)}^2. \quad (5.14)$$

NEES je ve své podstatě kvadrát vzdálenosti mezi odhadem a skutečností v rámci $\mathbf{P}_k^{-1}(i)$ normy.

Pro odstranění závislosti na dimenzi n_x a pro snížení neurčitosti byl definován [5, 24] průměrný NEES (*Average NEES*, ANEES)

$$\bar{\epsilon}_k = \frac{1}{n_x M} \sum_{i=1}^M \epsilon_k(i). \quad (5.15)$$

Pak při nestranném odhadu a předpokladu gaussovského rozdělení chyby odhadu, tj. $\tilde{\mathbf{x}}_k \sim \mathcal{N}(0, \boldsymbol{\Sigma}_k)$, je střední hodnota ANEES vyjádřena [24] vztahem

$$\text{E}[\bar{\epsilon}_k] = \frac{\text{tr}(\mathbf{P}_k^{-1} \boldsymbol{\Sigma}_k)}{n_x}, \quad (5.16)$$

z kterého jasně plyne závislost ANEES na matici MSE. Dále ze vztahu (5.16) vyplývá jasný požadavek na ideální hodnotu ANEES, která by měla být co nejbližší k hodnotě 1.

V praxi se běžně používá χ^2 test věrohodnosti založený na ANEES [5, 24], nicméně toto řešení není dále doporučováno. Důvodem tohoto postupného upouštění od jinak zavedeného postupu především v oblasti sledování cílů jsou některé nežádoucí vlastnosti testů založených na NEES. Tyto vlastnosti jsou podrobně diskutovány v [19, 24, 26] a patří mezi ně například:

- pesimističnost ve vztahu k chybám,
- schopnost pouze vyvrátit hypotézu o věrohodnosti, nikoliv vyvrátit nevěrohodnost
- přílišná závislost na aritmetickém průměru vlastních čísel, daná (5.16) (MSE matice a kovarianční matice mohou být velmi odlišné, i když bude průměr vlastních čísel věrohodnostní matice $\mathbf{P}_k^{-1}\boldsymbol{\Sigma}_k$ roven 1, viz numerické ilustrace v [26]).

Řešení těchto nedostatků jsou v zásadě dvě. Buď testovat hypotézy pomocí upravených testů prezentovaných v [26]. Nebo se spíše soustředit na číselné vyhodnocení (ne)věrohodnosti, k čemuž je nejvhodnější následující metrika, tzv. index nevěrohodnosti.

5.2.3 Index nevěrohodnosti a indikátor inklinace - NCI a I^2

V současné době se jako nejlepší nástroj pro kvantifikaci věrohodnosti estimátorů jeví index nevěrohodnosti (*Noncredibility Index*, NCI). Tato metrika vychází z úvahy, že k vyšetření rozdílu mezi \mathbf{P}_k a $\boldsymbol{\Sigma}_k$ můžou stejně dobře posloužit i maticové inverze \mathbf{P}_k^{-1} a $\boldsymbol{\Sigma}_k^{-1}$. NCI porovnává NEES estimátoru $\epsilon_k(i) = \tilde{\mathbf{x}}_k^T(i)\mathbf{P}_k^{-1}(i)\tilde{\mathbf{x}}_k(i)$ s NEES dokonale věrohodného estimátoru $\epsilon_k^*(i) = \tilde{\mathbf{x}}_k^T(i)\boldsymbol{\Sigma}_k^{-1}\tilde{\mathbf{x}}_k(i)$ poměrem těchto dvou normovaných kvadrátů chyby. K odstranění neurčitosti je použit geometrický průměr, který má nejvhodnější vlastnosti pro průměrování poměrů. Výsledný vztah obsahuje navíc ještě zesílení 10 a z výpočetních důvodů také logaritmus, jak tomu bývá v obdobných vzorcích:

$$\text{NCI} = \frac{10}{M} \sum_{i=1}^M \left| \log_{10} \frac{\epsilon_k(i)}{\epsilon_k^*(i)} \right|. \quad (5.17)$$

Dále je definován indikátor inklinace (*Inclination Indicator*, I^2) opomenutím absolutní hodnoty z (5.17)

$$I^2 = \frac{10}{M} \sum_{i=1}^M \log_{10} \frac{\epsilon_k(i)}{\epsilon_k^*(i)}. \quad (5.18)$$

Ve vztahu (5.17) je NCI definováno jako aritmetický průměr absolutních hodnot logaritmů. Není bez zajímavosti, že v pionýrských pracích ohledně indexu nevěrohodnosti bylo NCI definováno vztahem (5.18), který již autor v článku [22] nazývá indikátorem inklinace. Proto se lze setkat s různými výklady NCI [33, 6].

Důležitější je ale prozkoumat, proč byla do vztahu (5.17) vložena absolutní hodnota. Pokud bude estimátor v každé jednotlivé iteraci simulace optimistický ($\epsilon_k(i)/\epsilon_k^*(i) > 1$) nebo pesimistický ($\epsilon_k(i)/\epsilon_k^*(i) < 1$), pak absolutní hodnota není potřeba a $\text{NCI}=I^2$. Jiná situace ovšem nastane, pokud poměr $\epsilon_k(i)/\epsilon_k^*(i)$ bude někdy pesimistický a někdy optimistický (bude kolísat kolem hodnoty 1). Tyto poměry se navzájem vyruší a není těžké si představit situaci, kdy bude estimátor silně nevěrohodný a přitom indikátor inklinace bude potvrzovat věrohodnost.

Vzhledem k výše uvedeným skutečnostem je dle [22] žádoucí používat NCI a I^2 současně. Nejprve indexem NCI určit úroveň věrohodnosti (čím blíže k hodnotě 0, tím

lépe) a následně indikátorem I^2 přezkoumat, zda estimátor spíše tíhne k optimismu (větší než 0) či k pesimismu (menší než 0).

Vztahy (5.17) a (5.18) platí pro nestranné odhady, řešení pro stranné odhady lze nalézt v [22].

Index nevěrohodnosti způsobil ve světě hodnocení kvality odhadu malou revoluci, protože konečně umožnil adekvátně zhodnotit kovarianční matice a jejich kvalitu. Důvodem je, že NCI a I^2 společně mají všechny hezké vlastnosti, které si lze u hodnocení věrohodnosti přát: vyváženost, relativní vyjádření, dimenzionálně invariantní a jasně vytyčenou optimistickou/pesimistickou definitnost. Ostatně absence NCI v české literatuře je i jedním z důvodů vzniku této diplomové práce.

5.3 Cramér-Raova mez - porovnání MSE s mezí poznatelnosti

Zajímavým způsobem hodnocení kvality odhadu stavu je porovnání odhadů dodaných estimátorem s objektivním limitem poznatelnosti stavu. Tomuto limitu se říká Cramér-Raova (CR) mez nebo také Cramér-Raova dolní mez. Tato mez původně sloužila k omezení variance chyby odhadu zdola při úloze odhadu konstantního parametru, ale později se ukázalo [38], že lze CR mez s úspěchem použít i v úloze odhadu náhodné proměnné. Odtud byl jen krůček k aplikaci meze v úloze odhadu stavu. To vyústilo k odvození aposteriorní CR meze [15]. V této práci bude ukázána aplikace CR meze pro rekurzivní filtraci při systému s gaussovskými šумы.

Pro zavedení CR meze bude nejprve uveden případ odhadu náhodné proměnné \mathbf{x} za předpokladu známé sdružené hustoty pravděpodobnosti $p(\mathbf{x}, \mathbf{z})$. Matice střední kvadratické chyby (MSE) odhadu $\hat{\mathbf{x}}$ definovaná jako

$$\Sigma(\hat{\mathbf{x}}) \triangleq \mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T] \quad (5.19)$$

je zdola ohraničena nerovností

$$\Sigma(\hat{\mathbf{x}}) \geq \mathbf{C}(\mathbf{x}), \quad (5.20)$$

kde matice $\mathbf{C}(\mathbf{x})$ je Cramér-Raova mez definovaná jako inverze Fisherovy informační matice $\mathbf{J}(\mathbf{x})$, tj.

$$\mathbf{C}(\mathbf{x}) = \mathbf{J}^{-1}(\mathbf{x}), \quad (5.21)$$

$$\mathbf{J}(\mathbf{x}) \triangleq -\mathbb{E} \left[\frac{\partial}{\partial \mathbf{x}} \left[\frac{\partial}{\partial \mathbf{x}} \ln p(\mathbf{x}, \mathbf{z}) \right]^T \right]. \quad (5.22)$$

Níže jsou uvedeny konečné vztahy definující rekurzivní výpočet Cramér-Raovy meze pro systém s gaussovskými šумы (2.1). Kompletní odvození CR meze lze nalézt v [35, 37]. Filtrační a jednokroková prediktivní CR mez jsou dány vztahy

$$\mathbf{C}_{k|k} = \mathbf{C}_{k|k-1} - \mathbf{C}_{k|k-1} \mathcal{H}_k^T (\mathcal{H}_k \mathbf{C}_{k|k-1} \mathcal{H}_k^T + \mathbf{R}_k)^{-1} \mathcal{H}_k \mathbf{C}_{k|k-1}, \quad (5.23)$$

$$\mathbf{C}_{k+1|k} = \mathcal{F}_k \mathbf{C}_{k|k} \mathcal{F}_k^T + \mathbf{Q}_k, \quad (5.24)$$

$$(5.25)$$

kde jsou Jacobiho matice

$$\mathcal{F}_k = \left. \frac{\partial \mathbf{f}_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_k}, \quad (5.26)$$

$$\mathcal{H}_k = \left. \frac{\partial \mathbf{h}_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_k} \quad (5.27)$$

funkcemi skutečného stavu \mathbf{x}_k a počáteční podmínka je $\mathbf{C}_{0|-1} = \mathbf{P}_{0|-1}$.

Rovnice definující rekurzivní výpočet CR meze jsou formálně shodné s rovnicemi kovariancí $\mathbf{P}_{k|k}$ a $\mathbf{P}_{k+1|k}$ rozšířeného Kalmanova filtru (EKF). Jediným rozdílem je využití Jacobiho matic. Zatímco v EKF je využití Jacobiho matic důsledkem linearizace v bodech $\hat{\mathbf{x}}_{k|k}$ a $\hat{\mathbf{x}}_{k+1|k}$, v případě CR meze vychází Jacobiho matice z definice Fisherovy informační matice a její závislosti na skutečném stavu \mathbf{x}_k . Pokud by uvažovaný systém (2.1) byl lineární, tj. $\mathbf{f}_k(\mathbf{x}_k) = \mathbf{F}_k \mathbf{x}_k$, $\mathbf{h}_k(\mathbf{x}_k) = \mathbf{H}_k \mathbf{x}_k$, pak by rekurzivní rovnice CR meze vůbec nezávisely na skutečném stavu \mathbf{x}_k a byly by identické se vztahy pro kovarianční matici Kalmanova filtru.

Teoreticky identičnost matic MSE a CR meze, tj. rovnost v (5.20), nastává pouze za předpokladů:

- použití optimálního estimátoru s návrhovým kritériem MSE, tj. minimalizace ve smyslu podmíněné střední hodnoty
- shodnosti počátečních podmínek estimátoru a skutečných počátečních podmínek systému nebo alespoň pokud je systém v rovnovážném stavu.

Bohužel tyto podmínky v praxi nikdy nelze splnit už jen proto, že v případě nelineárních systémů se jedná vždy o aproximativní řešení odhadu stavu. Nicméně CR mez se hojně využívá ke zjištění, jak velké chyby aproximací jsou a jak moc by se odhad mohl přiblížit k mezi poznatelnosti. Porovnání většinou vychází z hodnocení rozdílu těchto dvou matic za pomoci některé z metrik představených v sekci 5.1.

Kapitola 6

Metodika hodnocení kvality odhadu

V této kapitole bude představena metodika hodnocení kvality odhadu stavu dynamických systémů na základě metrik představených v této práci.

Základem zevrubné offline analýzy kvality odhadu je dostatečný počet M Monte Carlo simulací, protože každá metrika je odhadem skutečné hodnoty, a u většiny metrik se zvyšuje přesnost těchto odhadů úměrně s \sqrt{M} . Metriky jsou koncipovány tak, že lze vyhodnocovat kvalitu odhadu pro každý časový krok k zvlášť jediným číslem. Proto je asi nejlepší praxí pro každou metriku hodnotit křivku závislosti hodnoty metriky na časovém kroku k . Pokud to není žádoucí, a skutečným cílem je získat pouze jediné číslo přes všechny iterace MC i časové kroky k , lze výsledky metrik pro časové okamžiky k nějakým způsobem zprůměrovat. K tomuto zprůměrování lze v podstatě použít jakoukoliv absolutní metriku ze sekce 4.1.

Nyní již k samotné metodice hodnocení kvality odhadu. Prvním krokem je jistě jakkoliv vyhodnotit chybu bodového odhadu stavu \mathbf{x}_k . K tomu lze využít kteroukoliv metriku ze sekce 4.1, záleží na daném problému a preferencích uživatele. K usnadnění výběru může pomoci souhrnná tabulka¹ 6.1. Absolutní metriky chyby odhadu jsou většinou první volbou oproti relativním metrikám, které se hodí spíše pro vzájemné porovnání různých estimátorů v rámci různých řešeních úloh.

Tabulka 6.1: Metriky absolutní chyby odhadu a jejich relativní verze

Metrika	Zařazení	Výhody	Nevýhody	Vhodné pro
RMSE, RMSRE	pesimistické	populární, silná vazba s návrhovým kritériem MSE a standardní odchylkou	dominance velkých chyb, nejasná fyzikální interpretace	pravděpodobnostní analýzu
AEE, ARE		jasná fyzikální interpretace	protežování velkých chyb	měření průměrné chyby
IMRE		velmi robustní	iterativní výpočet	hledání centrální tendence
GAE, GRE	vyvážené	robustní s ohledem na extrémní chyby	nepopulární	průměrování poměrů
HAE, HRE	optimistické	robustní s ohledem na velké chyby	dominance malých chyb, nepopulární, nejasná fyzikální interpretace	průměrování poměrů s konstantním základem, ověření správného chování
medián	vyvážené	robustní vůči extrémním chybám	těžké vypočítat a matematicky uchopit	hledání typické chyby
modus				

¹tabulky uvedené v této kapitole čerpají z tabulek v [21]

Pro právě takové porovnání estimátorů jsou neocenitelné alternativní metriky relativní chyby odhadu ze sekce 4.2 shrnuté v tabulce 6.2. Tyto metriky mají schopnost srovnáním filtračních a prediktivních odhadů či vyhodnocením přínosu příchozích měření prověřit správnou vnitřní funkci estimátorů (kvalita filtračního a prediktivního kroku) a odhalit případná selhání (např. nesoulad mezi modelem a generátorem dat, špatná predikce či filtrace vinou chyby aproximací). Tyto skryté vady by při hodnocení standardními metrikami mohly zůstat opomenuty a později by se mohly negativně projevit při reálném nasazení.

Tabulka 6.2: Další metriky relativní chyby odhadu

Metrika	Zařazení	Výhody	Nevýhody	Vhodné pro
BEEQ	vyvážené	robustní vůči chybě predikce, jasná interpretace	závisí na chybě měření	porovnání kvality odhadu bayesovských estimátorů
EMER		robustní vůči chybě měření, jasná interpretace	závisí na chybě predikce	vyhodnocení zlepšení odhadu po příchodu nového měření
BERF		robustní vůči scénářím, hezké vlastnosti	nepopulární, nejasná fyzikální interpretace	porovnání kvality odhadu bayesovských estimátorů

V některých aplikacích není ani tak důležité jak kvalitní odhad je ve smyslu absolutní chyby odhadu, ale předmětem zkoumání je, zda je odhad v toleranci (např. při navádění rakety na cíl). Pokud tolerance není zřejmá ze zadání úlohy, je vhodné považovat toleranci za funkci měření (špatné měření nemůže poskytnout dobré odhady). K tomu byly definovány v sekci 4.3 míry četnosti shrnuté v tabulce 6.3.

Tabulka 6.3: Míry četnosti

Metrika	Zařazení	Výhody	Nevýhody	Vhodné pro
Míra úspěšnosti	optimistické	robustní vůči chybě měření, jasná interpretace	závisí na chybě predikce	aplikace typu hit-or-miss
Míra neúspěšnosti	pesimistické			měření špatné kvality

Dalším krokem hodnocení kvality odhadu obvykle bývá vyhodnocení věrohodnosti odhadu, tedy zda estimátor vytrvale nadhodnocuje či podhodnocuje kovarianční matici, která představuje vlastní hodnocení chyb estimátoru. K tomuto problému lze přistoupit dvěma způsoby. Buď testováním hypotéz o věrohodnosti odhadu nebo kvantifikací míry věrohodnosti. Obecně je preferován druhý způsob pomocí metrik věrohodnosti představených v sekci 5.2. Dvě nejdůležitější metriky jsou shrnuty do tabulky 6.4.

Tabulka 6.4: Metriky věrohodnosti

Metrika	Zařazení	Výhody	Nevýhody	Vhodné pro
ANEES	pesimistické	populární	nesymetrické, neporovnatelné výsledky pro různé estimátory	testování hypotézy věrohodnosti estimátoru (pouze zamítnutí)
NCI	vyvážené	symetrické, relativní, moderní	-	kvantifikaci věrohodnosti

Hodnotit kvalitu odhadu lze také z pohledu míry přiblížení se k objektivnímu limitu poznatelnosti systému - Cramér-Raově mezi. Výpočet CR meze pro dynamický systém je dán vztahy uvedenými v sekci 5.3. K výpočtu je potřeba znát skutečnou trajektorii

systemu. V případě MC simulace je těchto trajektorií více a je třeba výsledky nějakým způsobem zprůměrovat. To se většinou dělá průměrováním Jacobiho matic a následným jediným výpočtem CR meze.

Seřazení estimačních algoritmů dle vícero hledisek

Nezřídka se odborníci z oboru filtrace setkávají s problémem výběru vhodného filtračního algoritmu pro nějaký konkrétní problém. V této práci byly vyčteny významné metriky kvality odhadu (poskytovaného lokálními filtry), které lze pro tento výběr využít. Každá taková metrika je pouze jistým kompromisem a hodnotí pouze v jistých ohledech. Zvolením konkrétní metriky řešitel získá sice pořadí jím zvolených filtračních algoritmů, ale toto pořadí je na zvolené metrice závislé (tj. na zohlednění dílčích aspektů). Jistě je chválihodné zohlednit co nejvíc aspektů, a tudíž aplikovat více druhů hodnocení kvality. Otázkou ale zůstává, jak naložit s dílčími výsledky a jak určit celkové pořadí. Tato sekce dává návod, jak se s tímto problémem vypořádat využitím Pitmanovy míry blízkosti [30].

Pitmanova míra blízkosti (*Pitman's Closeness Measure*, PCM) je definována tak, že ze dvou odhadů δ_1 a δ_2 reálného parametru θ je lepší první odhad δ_1 pokud platí

$$P_{\theta}(|\delta_1 - \theta| < |\delta_2 - \theta|) > \frac{1}{2}. \quad (6.1)$$

Z této míry vychází míra vzájemných srovnání (*Multiple-Attribute Competition Measure*, MCM) navržená v [44], ze které je nakonec vypočten vektor ohodnocení estimatorů. Nejprve definujeme rozdíl mezi estimátory 1 a 2 v rámci metriky a_i jako

$$m_{1,2}(a_i) \triangleq \begin{cases} 1 & \text{pokud je lepší filtr 1 ve smyslu Pitmana,} \\ 0.5 & \text{pokud jsou si oba filtry rovny ve smyslu Pitmana,} \\ 0 & \text{pokud je lepší filtr 2 ve smyslu Pitmana,} \end{cases}$$

Pak je MCM definována

$$M_{1,2}(a) = \frac{1}{n} \sum_{i=1}^n m_{1,2}(a_i), \quad (6.2)$$

kde a je vektor všech použitých metrik (rozměru n). Z definice vyplývá, že platí vztah $M_{1,2}(a) + M_{2,1}(a) = 1$.

Nyní lze zkonstruovat matici MCM pro porovnání N estimatorů

$$X_{MCM} \triangleq \begin{bmatrix} M_{1,1}(a) & \dots & M_{1,N}(a) \\ \vdots & \ddots & \vdots \\ M_{N,1}(a) & \dots & M_{N,N}(a) \end{bmatrix}, \quad (6.3)$$

ze které již není obtížné získat vektor ohodnocení estimatorů \mathbf{r} . Jednou z metod je využít lineárního mapování. Pak nalezení vektoru ohodnocení přejde na nalezení vlastních vektorů matice MCM [44]. Perron-Frobeniův teorém [39] o pozitivních maticích (MCM

matice je pozitivní, pokud jsou případné nuly nahrazeny malými kladnými čísly) zaručuje existenci takového pozitivního vektoru náležícího nejvyššímu kladnému vlastnímu číslu. Tento vektor lze jednoduše vypočítat [27] limitou

$$\mathbf{r} = \lim_{n \rightarrow \infty} \frac{\mathbf{X}_{MCM}^n \mathbf{r}_0}{\|\mathbf{X}_{MCM}^n \mathbf{r}_0\|}, \quad (6.4)$$

kde \mathbf{r}_0 je libovolný kladný vektor, např. $[1, \dots, 1]^T$.

MCM je zajímavou možností pro řešení problému výběru vhodného estimátoru, protože umožňuje porovnat mnoho různých estimátorů navržených pro daný systém respektováním libovolného množství metrik společně. Výsledkem je vektor ohodnocení estimátorů dávající jasné pořadí.

Kapitola 7

Aplikace metrik ve vybraných úlohách

V nadcházejících odstavcích bude předvedena aplikace metrik na 3 typické úlohy. V části 7.1 bude předvedena základní úloha sledování - odhad polohy pohybujícího se objektu. V části 7.2 bude uvedena úloha z oblasti sledování chemických procesů. Na závěr, v části 7.3 bude použit jednorozměrný růstový model, používaný v ekonometrii.

7.1 Odhad polohy pohybujícího se objektu

V této části bude pozornost upřena na hodnocení kvality v úloze odhadu polohy pohybujícího se objektu. Jedná se o jednu ze základních úloh sledování a navigace. Tento konkrétní příklad je z knihy [7] a je hojně používán pro testování estimátorů a jejich schopnost vypořádat se s nelinearitou měření při lineární dynamice.

V úloze je sledován pohyb tankeru, který vplouvá do rušného přístavu podle přesných instrukcí. Tyto instrukce byly dodány řídicím střediskem přístavu, a pro bezpečnost dopravy je bezpodmínečně nutné se jimi řídit. Cílem úlohy je ověřit, zda loď dodržuje předepsanou trajektorii trasy pomocí měření prováděného přízemním radarem dodávajícím úhel a vzdálenost.

V úloze je uvažována stavová reprezentace vycházející ze systému (2.1) s modelem konstantní rychlosti ve 2D. Vektor stavu je definován fyzikálními proměnnými polohy a rychlosti v osách x a y , tj. $\mathbf{x} := [x^x, x^y, \dot{x}^x, \dot{x}^y]^T$. Vektor inkrementálního řízení je definován $\mathbf{u} := [-\Delta\dot{x}^{x0} - \Delta\dot{x}^{y0}]^T$. Celý systém je dán rovnicí dynamiky modelu

$$\mathbf{x}_{k+1} = \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{u}_k + \mathbf{w}_k \quad (7.1)$$

a rovnicí měření

$$\mathbf{z}_k = \begin{bmatrix} \arctan \frac{x_k^y}{x_k^x} \\ \sqrt{(x_k^x)^2 + (x_k^y)^2} \end{bmatrix} + \mathbf{v}_k. \quad (7.2)$$

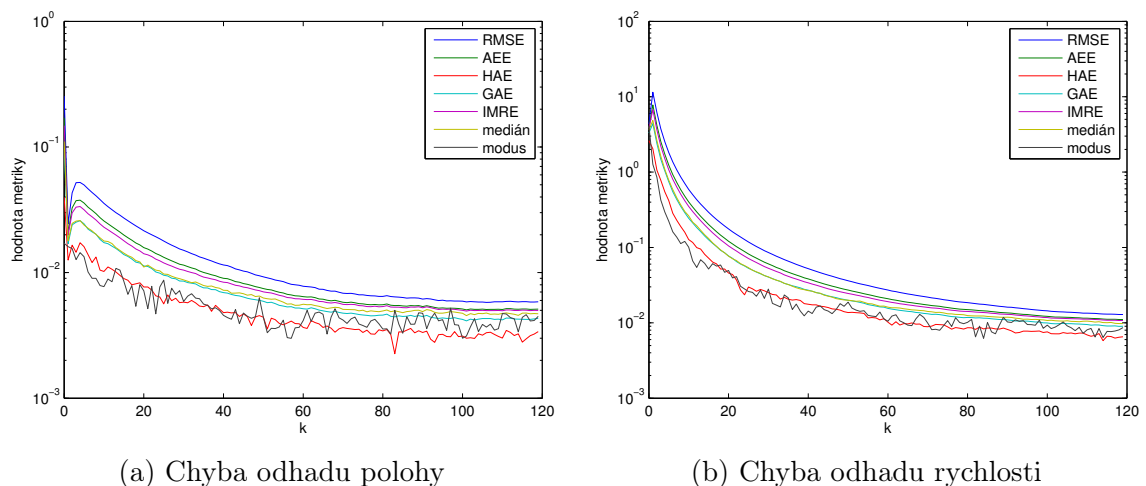
Systém byl simulován v prostředí MATLAB pomocí toolboxu nelineární filtrace (NEF) vyvinutým na katedře kybernetiky na ZČU v Plzni. V testovacím scénáři byly použity stejné parametry jako v [7], pouze doba simulace byla zkrácena a kovariance počátečního odhadu byla zvýšena pro lepší ilustraci vlastností metrik. Parametry tedy byly

následující:

počáteční podmínka	$\mathbf{x}_0 = [20 \text{ nm}, 50 \text{ nm}, 0 \text{ k}, -12 \text{ k}]^T$
	$\mathbf{P}_0 = \text{diag}(10, 10, 10, 10)$
kovarianční matice stavového šumu	$\mathbf{Q}_k = \text{diag}(10^{-6}, 10^{-6}, 10^{-6}, 10^{-6})$
kovarianční matice šumu měření	$\mathbf{R}_k = \text{diag}(4 \times 10^{-4}, 1 \times 10^{-4})$
frekvence měření	$\Delta T = 0,02 \text{ h}$
časových kroků simulace	$k_{max} = 120$
inkrementální řízení	$\mathbf{u}_{50} = [4,8 \text{ k}, 1,5 \text{ k}]^T$
počet nezávislých simulací	MC = 1000

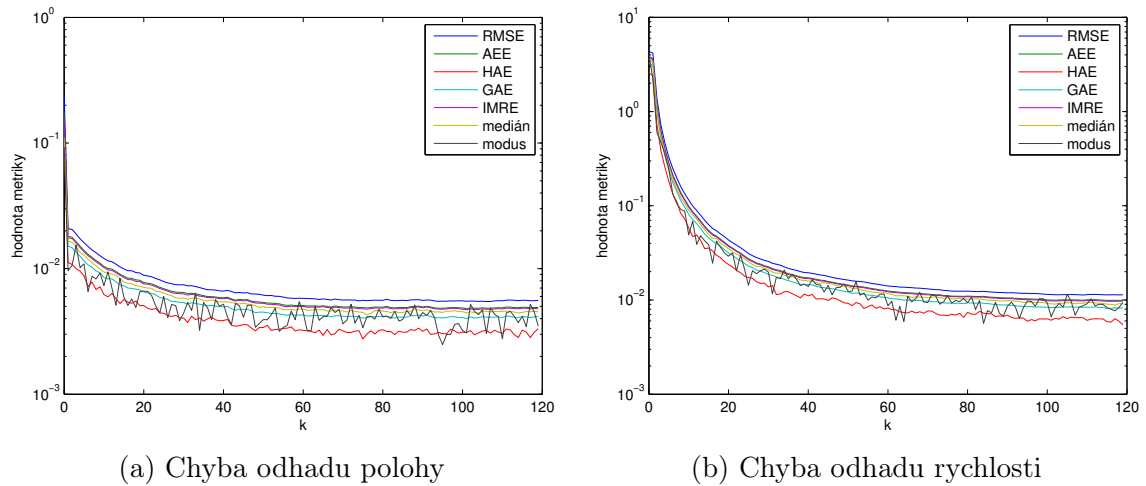
Výsledky

Tato úloha je typickou vzorovou ukázkou problému, kde je vhodné využít rozdělení stavu na fyzikálně shodné části, jak je diskutováno v sekci 4.4. Na obrázcích 7.1 a 7.2 jsou zobrazeny průběhy hodnot jednotlivých metrik absolutní chyby odhadu v čase pro EKF, resp. UKF. Z obrázků jsou výborně patrné vlastnosti těchto metrik. Především se potvrzuje nerovnost $HAE \leq GAE \leq AEE \leq RMSE$. Nejrozkošnější metrikou je podle očekávání modus, protože možných hodnot chyb je nekonečně mnoho, a modus se tak musí určit hledáním vrcholu vyhlazeného histogramu, což není triviální úloha. Za všimnutí rovněž stojí srovnání průběhu nejpesimističtější a nejoptimističtější metriky - RMSE a HAE. Konkrétně z obrázku 7.1b znázorňujícím chybu odhadu rychlosti tankeru poskytovaného EKF je patrné, že HAE je od začátku klesající funkcí oproti RMSE, které v kroku 1 dosáhne svého maxima a pak rovněž klesá. Z tohoto chování lze vyvozovat, že odhad rychlosti selhal v pár simulacích. Je to důsledkem nelinearity funkce měření a aproximačních chyb linearizace EKF.

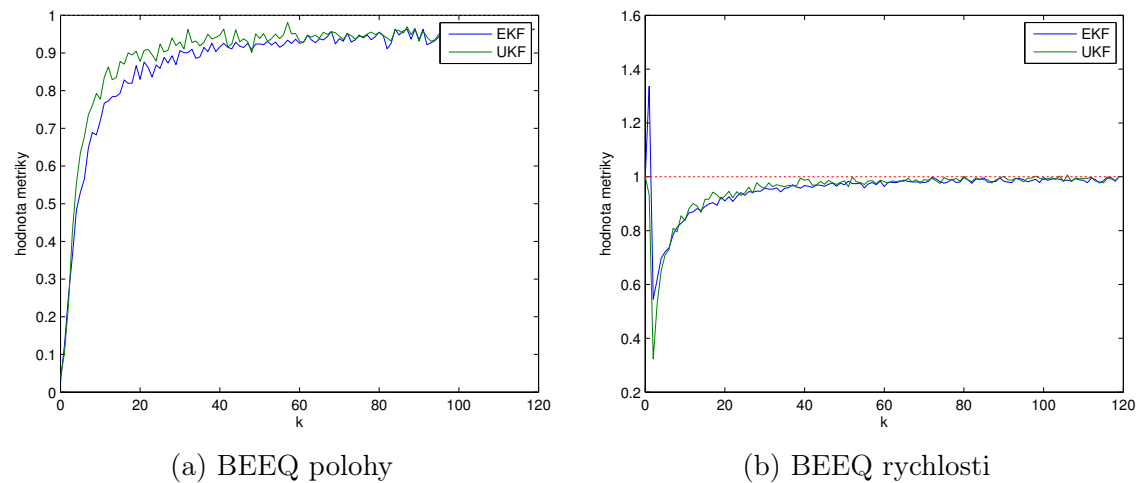


Obrázek 7.1: Porovnání průběhu hodnot metrik absolutní chyby odhadu EKF v čase.

Další metrikou je BEEQ - poměr chyby odhadu mezi filtračním a prediktivním krokem. Z obrázku 7.3a lze vyčíst, že se algoritmus UKF lépe a dříve vypořádá s neurčitostí než algoritmus EKF. Po nějaké době se hodnoty BEEQ obou estimátorů srovnají, protože je dynamika systému lineární. Na obrázku 7.3b je krásně vidět, že EKF při odhadu rychlosti v prvním kroku selhává, neboť BEEQ je výrazně vyšší než 1 (filtrační odhad je horší než prediktivní), což je důsledkem chyb linearizace funkce měření.

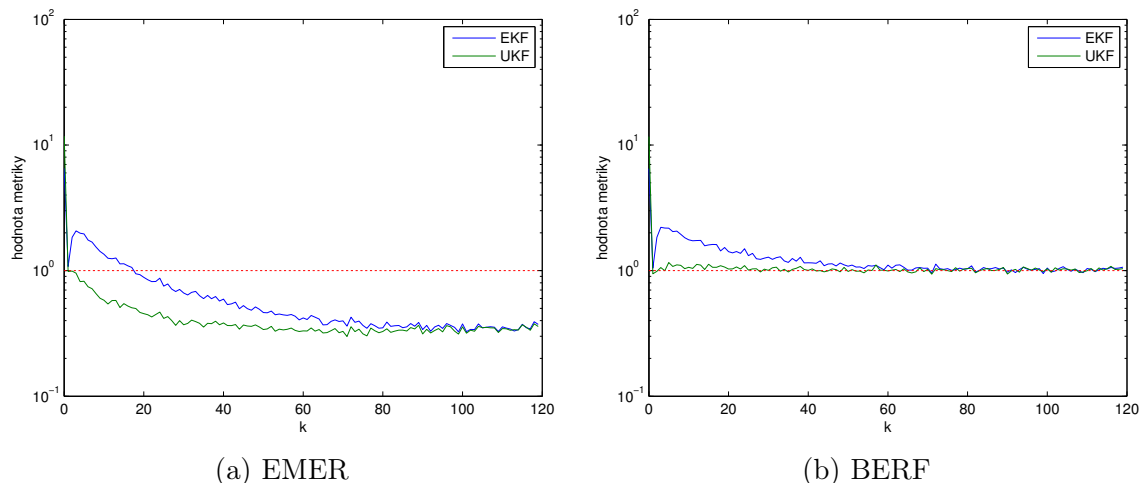


Obrázek 7.2: Porovnání průběhu hodnot metrik absolutní chyby odhadu UKF v čase.



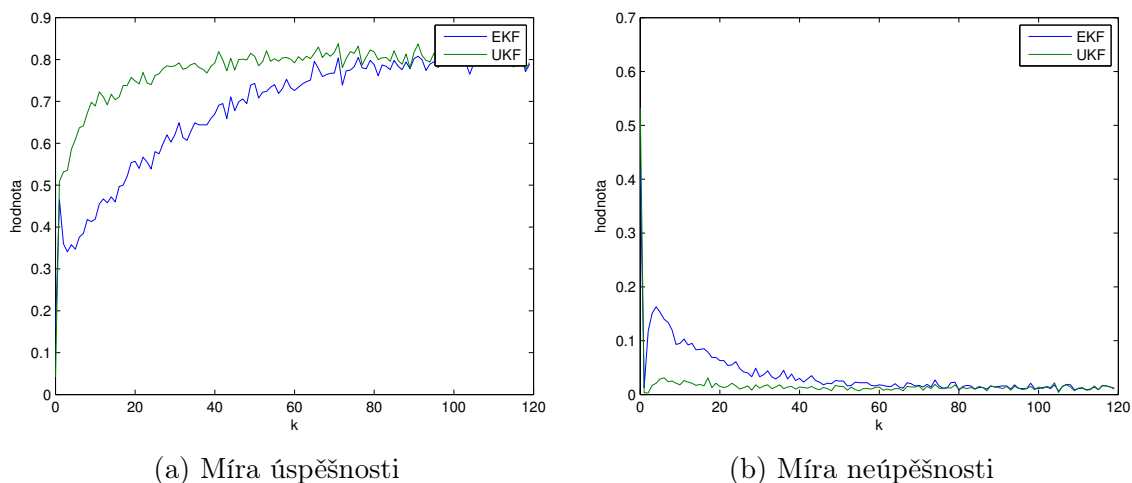
Obrázek 7.3: Průběh relativní metriky BEEQ v čase.

Pro zbylé metriky relativní chyby odhadu nelze rozdělit stav na fyzikálně si odpovídající části, proto jsou jejich výsledky o něco hůře interpretovatelné. Porovnání chyby filtračního odhadu s chybou triviálního jednokrokového odhadu určeného pouze z aktuálního měření (tedy tedy s chybou odhadu ve smyslu maximální věrohodnosti ignorujícího dynamiku) nabízí metrika EMER. Výsledné hodnoty této metriky jsou na obrázku 7.4a. Opět lze vyčíst, že odhad EKF byl zanesen daleko od skutečných hodnot a estimátor chvíli trvá, než se vrátí ke správným hodnotám. Oproti tomu z vývoje EMER pro UKF je vidět, jak estimátor postupně zpřesňuje svůj odhad a nenechává se vychýlit nepřesnými měřeními, naopak díky nim odhad ještě zpřesňuje. Na vedlejším obrázku 7.4b je znázorněn vývoj hodnot metriky BERF, která je kompromisem mezi BEEQ a EMER. Bohužel tato metrika nemá jasnou fyzikální interpretaci, a tak je třeba se spokojit s tvrzením jejího autora, že by hodnota dobrého estimátoru neměla překročit hodnotu 1. To se ani pro EKF ani pro UKF trvale nedaří, nicméně UKF se alespoň této hodnoty drží na rozdíl od EKF, které opět na začátku zaspí a pak pomalu konverguje k hodnotě UKF.



Obrázek 7.4: Průběh relativních metrik EMER a BERF v čase.

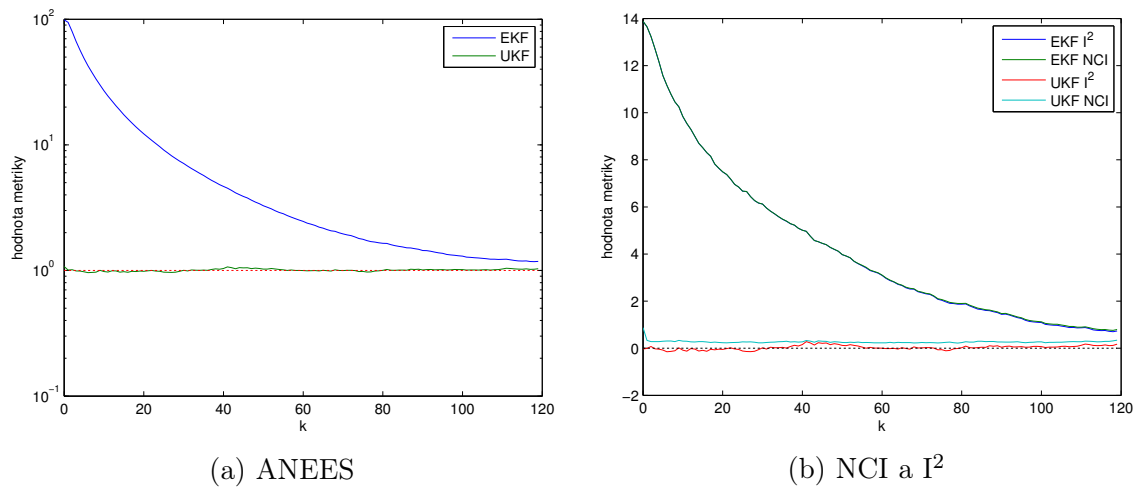
S metrikou EMER úzce souvisí metriky míra úspěšnosti a míra neúspěšnosti, protože oblasti ve kterých odhad přijmeme či zamítneme jsou definovány právě na základě EMER. Tyto dvě metriky jsou předmětem obrázků 7.5. Pro míru úspěšnosti byla zvolena taková oblast ($\sigma_s = 1$), ve které jsou odhady estimátoru lepší než triviální odhad zmíněný v předchozím odstavci. Pro míru neúspěšnosti byla zvolena taková oblast, ve které je odhad estimátoru $10\times$ horší než triviální odhad ($\sigma_f = 10$). Vlivem ojediněle přicházejících velmi přesných měření nelze očekávat ideální výsledky (míra úspěšnosti 100 %, míra neúspěšnosti 0 %), ale opět se jasně potvrzuje, že se UKF lépe než EKF vypořádává s nelinearitami měření.



Obrázek 7.5: Průběh relativních metrik míry úspěšnosti a míry neúspěšnosti v čase.

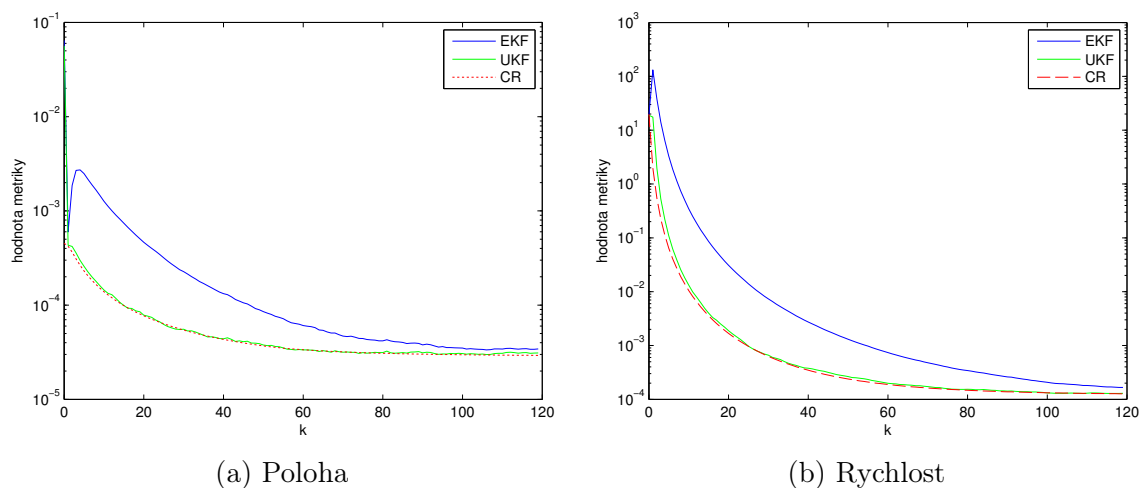
Jedním z novějších přístupů k hodnocení kvality odhadu, které se snaží tato práce reflektovat, je důraz na věrohodnost odhadu. Tento přístup vychází z úvahy, že minimální chyba odhadu není jediným cílem, ale podobně důležitá je věrohodnost odhadu reprezentovaná kovarianční maticí dodanou estimátorem. V práci byly uvedeny dvě metriky kvantifikující věrohodnost - ANEES a NCI (vždy dohromady s I^2). Jejich průběhy jsou zachyceny na obrázku 7.6. Z obou metrik vychází odhady UKF jako velmi

věrohodné (kovarianční matice odhadu stavu odpovídá MSE matici). V případě odhadů EKF je opět patrná pomalá konvergence po počátečním selhání.



Obrázek 7.6: Průběh metrik věrohodnosti v čase.

Poslední předvedenou metrikou je stopa Cramér-Raovy meze zobrazená na obrázku 7.7. CR mez je objektivním limitem poznatelnosti systému, a jako taková je vždy menší než MSE matice. Na obrázku 7.7a je vyobrazena stopa CR meze pro polohu a pro srovnání taktéž stopy polohové submatice MSE. Na obrázku 7.7b jsou tytéž stopy tentokrát pro rychlostní submatice. U EKF jsou opět patrné výše diskutované jevy s postupnou opožděnou konvergencí. O něco zajímavější je porovnání v případě UKF. Z grafů plyne, že odhady poskytované UKF se blíží mezi poznatelnosti. Může se zdát, že je UKF v některých časových okamžicích lepší než CR mez, ale ve skutečnosti tomu tak není. Jedná se totiž pouze o odhady (ať již CR meze nebo MSE matice) a také vzhledem k hodnotám kolem 10^{-4} je třeba brát do úvahy numerické chyby.



Obrázek 7.7: Průběh stopy MSE matice a CR meze v čase

Obecně lze k tomuto příkladu říci, že vzhledem ke znalosti všech parametrů je rozdíl mezi EKF a UKF patrný pouze v prvních krocích simulace, kdy je kovariance

(nejistota) vysoká. Jakmile se odhad dostatečně přiblíží skutečnosti, není rozdíl mezi kvalitou odhadu EKF a UKF. Dynamika systému je totiž lineární, a tak zde není prostor pro chyby linearizace, protože se neprovádí. Každá metrika odhalí jiné módy a ve výsledku nám tyto metriky dají velmi pěkný přehled o tom, co se děje při estimaci a jaké chyby vznikají, a proč je tomu tak. Tento příklad byl rozebrán velmi podrobně, aby byly metriky komplexně představeny a čtenář získal lepší představu, k čemu jsou vhodné. V dalších dvou příkladech se některé výsledky podobají výsledkům z tohoto příkladu, proto budou uvedeny pouze ty grafy a metriky, ve kterých se objevují jevy hodné povšimnutí.

7.2 Odhad veličiny chemického procesu

Jako druhý ukázkový příklad byla zvolena reakce dvou plynů A a B ve směšovací reaktoru. Tento příklad je často používán jako testovací pro hodnocení kvalit estimátorů [11, 17]. Důvodem je silně nelineární dynamika systému, nepřesné měření a velmi nepřesná počáteční podmínka estimátoru. Na estimátor číhá nebezpečí, protože důsledkem aproximací se může dostat odhadu stavu dílčích tlaků do záporných hodnot, které jsou fyzikálně nemožné.

Stav systému je reprezentován dílčími tlaky plynů P_A a P_B , tj. $\mathbf{x} = [P_A, P_B]^T = [x_1, x_2]^T$. Chování plynů v reaktoru je dáno diskretizovanou rovnicí dynamiky

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} \frac{x_{1,k}}{1+2k_r\Delta t} \\ x_{2,k} + \frac{k_r\Delta t x_{1,k}}{1+2k_r\Delta t} \end{bmatrix} + \mathbf{w}_k, \quad (7.3)$$

a měření je pouze celkový tlak v reaktoru, tj.

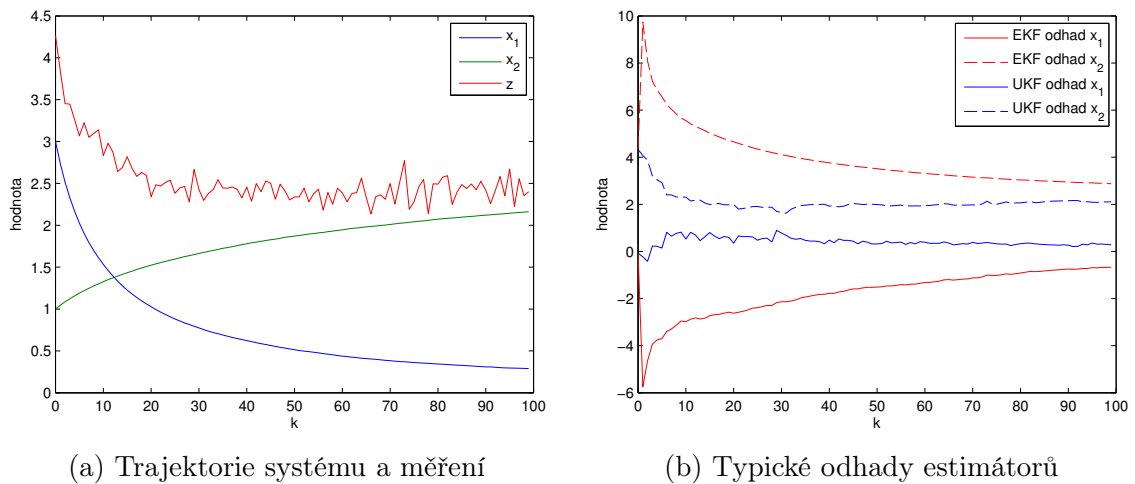
$$\mathbf{y}_k = [1, 1]\mathbf{x}_k + \mathbf{v}_k. \quad (7.4)$$

Parametry použité pro simulaci jsou následující:

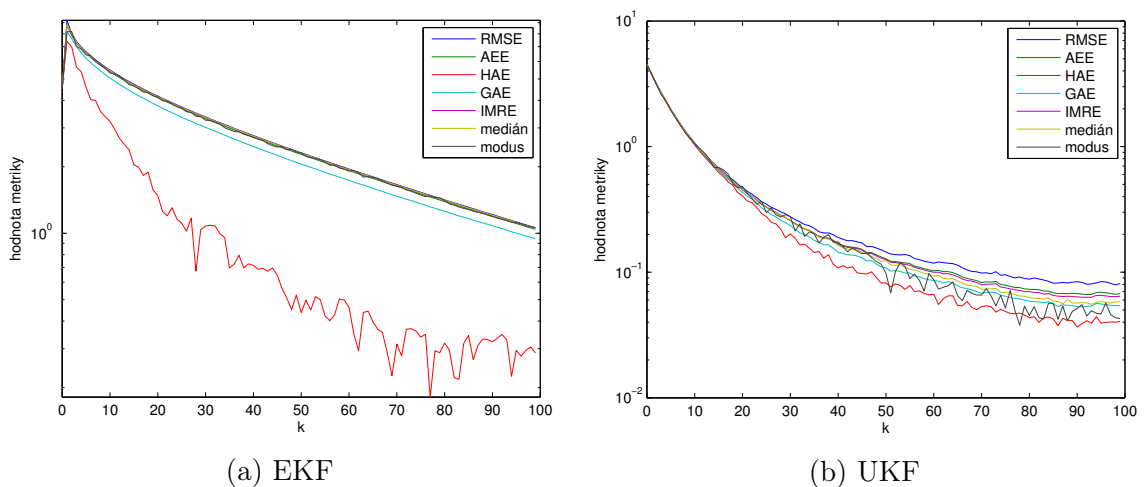
počáteční podmínka generátoru dat	$\mathbf{x}_0 = [3, 1]^T$
počáteční podmínka pro estimaci	$\hat{\mathbf{x}}_0 = [0, 1, 4, 5]^T$ a $\mathbf{P}_0 = \text{diag}(36, 36)$
kovarianční matice stavového šumu	$\mathbf{Q}_k = \text{diag}(1 \times 10^{-6}, 1 \times 10^{-6})$
kovarianční matice šumu měření	$\mathbf{R}_k = 1 \times 10^{-2}$
frekvence měření	$\Delta t = 0,1$
reakční rychlost	$k_r = 0,16$
časových kroků simulace	$k_{max} = 100$
počet nezávislých simulací	MC = 1000

Výsledky

Vzhledem k tomu, že se systém díky nulové varianci počátečního stavu pohybuje stále po stejné trajektorii, může být tato trajektorie prezentována na obrázku 7.8a. Pro porovnání je součástí téhož obrázku vývoj měření (součet stavů a šumu) z jedné simulace, aby byla doložena náročnost úlohy (nepřesné měření). Vzhledem k nelinearitám a náročnosti úlohy odhady poskytované EKF většinou nekonvergují v časovém horizontu simulace oproti odhadům poskytovaným UKF, které vždy konvergují přijatelně rychle (viz obrázek 7.8b, kde jsou zobrazeny typické odhady obou estimátorů).



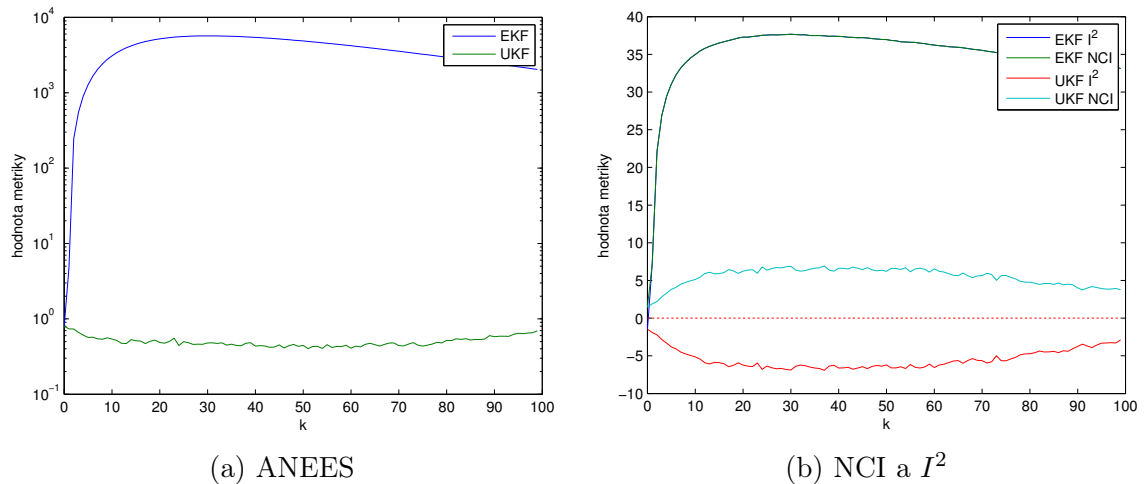
Metriky absolutní chyby odhadu jsou obsahem obrázků 7.9. V případě EKF je rozdíl mezi HAE a ostatními metrikami daleko markantnější než v minulém příkladu. Tentokrát se totiž EKF chovalo nežádoucím způsobem (viz obr. 7.8b) v 971 případech z 1000 simulací. Proto jsou všechny metriky s výjimkou optimistického HAE téměř za jedno. Právě tento rozdíl mezi HAE a ostatními metrikami lze interpretovat jako velmi špatné chování estimátoru v drtivé většině simulací. Hodnoty metrik pro UKF indikují typické dobré chování estimátoru.



Obrázek 7.9: Porovnání metrik absolutní chyby odhadu v čase.

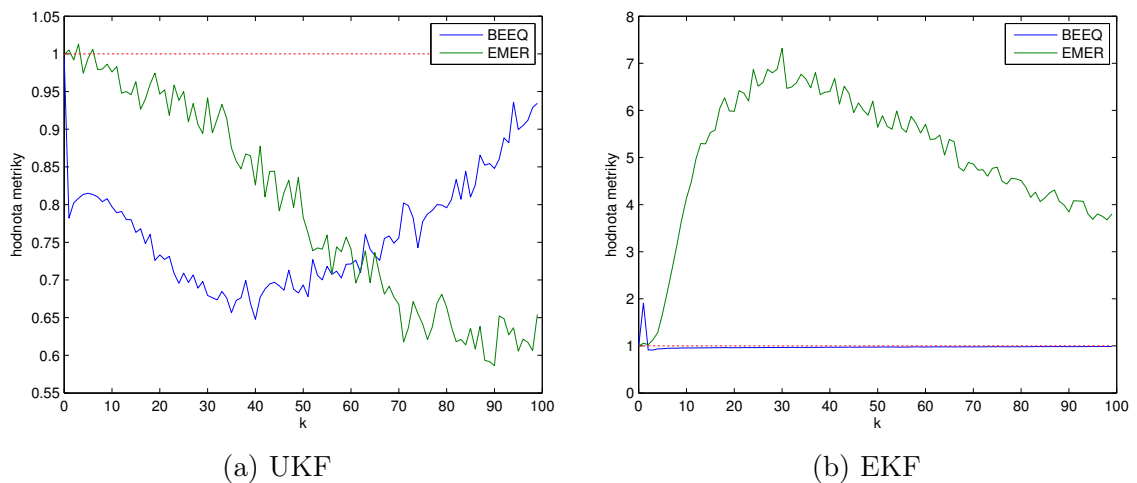
K objasnění pomalé konvergence odhadů produkovaných EKF pomohou grafy věrohodnosti vykreslené na obrázku 7.10. Z obou obrázků je patrné, že EKF nadhodnocuje kvalitu svých odhadů (malá kovarianční matice chyby odhadu). Z tohoto testu nevyšel ani UKF s čistým štítem. Tento estimátor trvale podhodnocuje své vlastní hodnocení přesnosti odhadu, což způsobuje upřednostňování vlivu měření na úkor modelu. Viditelným důsledkem je rozkolísanost odhadů patrná na obrázku 7.8b.

K ilustrování správné funkce metrik BEEQ a EMER poslouží obrázek 7.11a, kde jsou zachyceny vývoje těchto metrik pro odhady UKF. Metrika EMER neustále klesá, což je způsobeno neustálým zpřesňováním odhadu vůči odhadu ve smyslu maximální



Obrázek 7.10: Průběh metrik věrohodnosti v čase.

věrohodnosti, tj. ignorujícímu dynamiku. Metrika BEEQ nejprve klesá (nižší vliv predikce oproti filtraci) a po dosažení kritické úrovně kolem kroku 40 začíná stoupat, což vyjadřuje naopak přejímání iniciativy predikcí a upozadování příchozích nepřesných měření. Přesně takto má vypadat žádoucí chování estimátoru při nepřesných měřeních. Na druhou stranu obrázek 7.11b reflektuje typicky špatné chování estimátoru. Zprvu velké BEEQ značí zanesení odhadu pryč špatným filtračním krokem a pak již jen velmi malé změny správným směrem kvůli malé kovarianční matici. Z EMER je opět poznat zhoršování přesnosti odhadu až do určité meze, kdy se konečně podaří nasměrovat odhady správným směrem.



Obrázek 7.11: Průběh relativních metrik BEEQ a EMER v čase.

7.3 Nestacionární růstový model používaný v ekonometrii

Posledním příkladem v této práci je jednorozměrný nestacionární růstový model, který je velmi populární v ekonometrii a byl použit například v [16, 41]. Systém je popsán rovnicí dynamiky a rovnicí měření takto:

$$x_k = 0,5x_{k-1} + 25\frac{x_{k-1}}{1 + x_{k-1}^2} + 8\cos(1,2(k-1)) + w_k, \quad (7.5)$$

$$z_k = \frac{x_k^2}{20} + v_k. \quad (7.6)$$

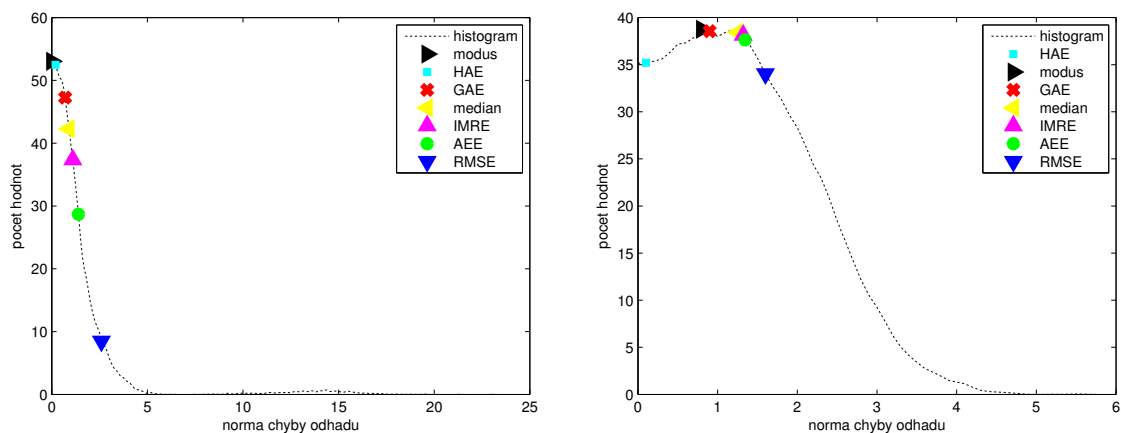
Specifikem tohoto systému je vysoký stupeň nelinearity dynamiky a bimodalita způsobená měřením obsahujícím druhou mocninu stavu. Tento příklad je tak náročný, že ani např. částicové filtry nevykazují uspokojivé chování.

Parametry použité pro simulaci jsou následující:

počáteční podmínka generátoru dat	$\mathbf{x}_0 = 0,1$
počáteční podmínka pro estimaci	$\hat{\mathbf{x}}_0 = 0,1$ a $\mathbf{P}_0 = 1$
kovarianční matice stavového šumu	$\mathbf{Q}_k = 1$
kovarianční matice šumu měření	$\mathbf{R}_k = 1$
časových kroků simulace	$k_{max} = 40$
počet nezávislých simulací	MC = 1000

Výsledky

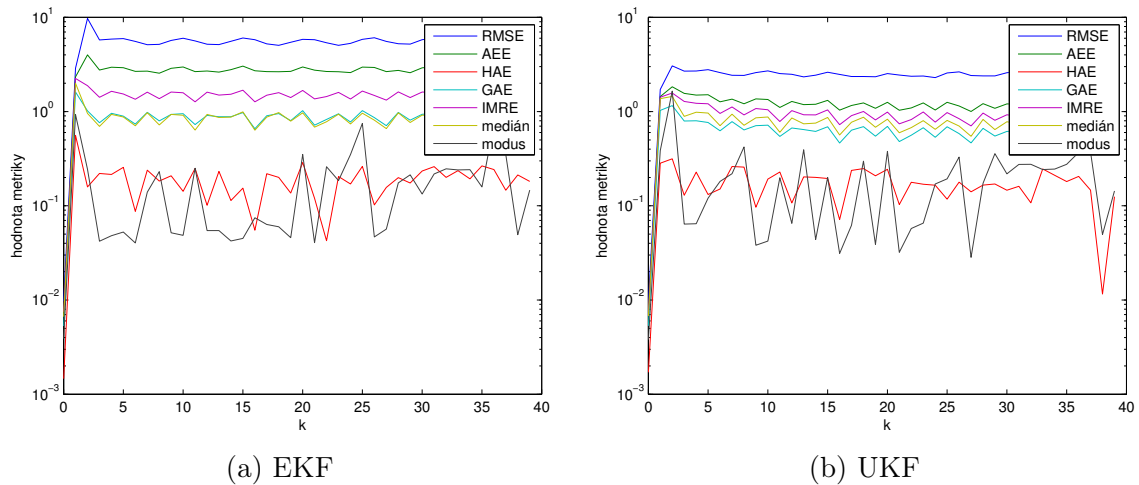
Než bude přistoupeno k samotnému vyhodnocení kvality odhadu pro tento příklad, bude využita náročnost příkladu k prezentování dvou různých histogramů normy chyby spolu s vypočtenými hodnotami metrik absolutní chyby. Tyto dva vyhlazené histogramy pocházejí z různých kroků simulace a jsou zobrazeny na obrázku 7.12. Cílem obrázku je ukázat, jak ovlivní histogram normy chyby výsledné hodnoty metrik. Je zřejmé, že HAE a RMSE se chovají skutečně extrémně optimisticky, resp. pesimisticky a při jejich nasazení je nutné zachovat maximální pozornost.



Obrázek 7.12: Histogramy normy chyby odhadu a hodnoty metrik.

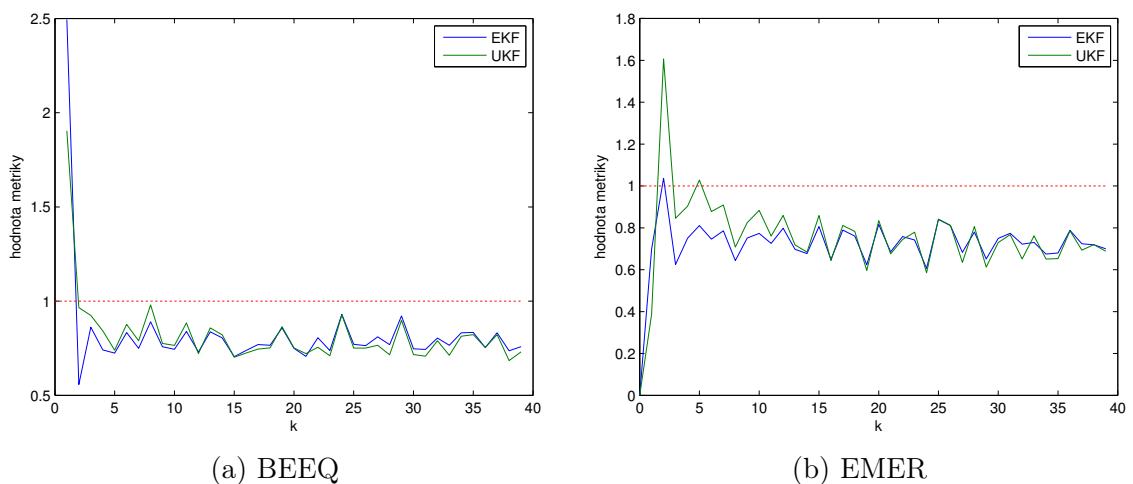
Nyní již k samotným výsledkům. Výsledky simulací jsou silně ovlivněny druhou mocninou stavu v měření. Pokud jsou hodnoty měření velmi malé, přichází ke slovu zmíněná bimodalita, což se bohužel stává téměř při každé simulaci.

Na obrázku 7.13 jsou již tradičně znázorněny průběhy jednotlivých metrik absolutní chyby odhadu. Výsledky UKF jsou opět o něco lepší než výsledky EKF, ale v obou případech se střídají malé a velké chyby, což způsobuje silně nelineární dynamika.



Obrázek 7.13: Průběh metrik absolutní chyby v čase.

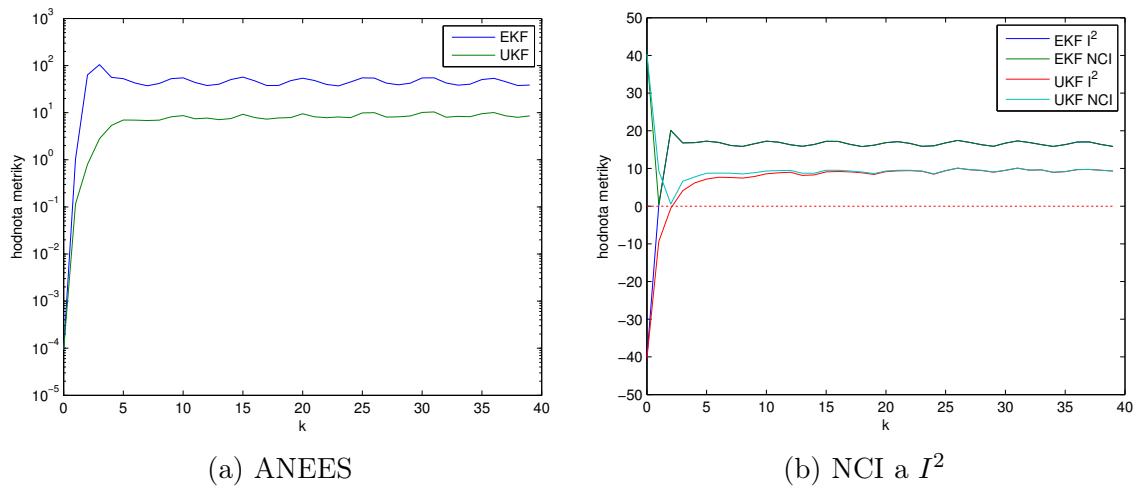
Metriky BEEQ a EMER (viz obr. 7.14) se ukázaly pro tento příklad jako málo vypovídající, neboť hodnotí oba estimátory zhruba stejně či dokonce UKF o něco hůře. Podstatné pro obě metriky je, že se jejich hodnoty pro estimátory drží téměř celý časový horizont simulace pod hodnotou 1. Vzhledem k počáteční podmínce má BEEQ v nultém kroku hodnotu ∞ a jen první krok naznačuje, že by UKF mohlo být lepší než EKF.



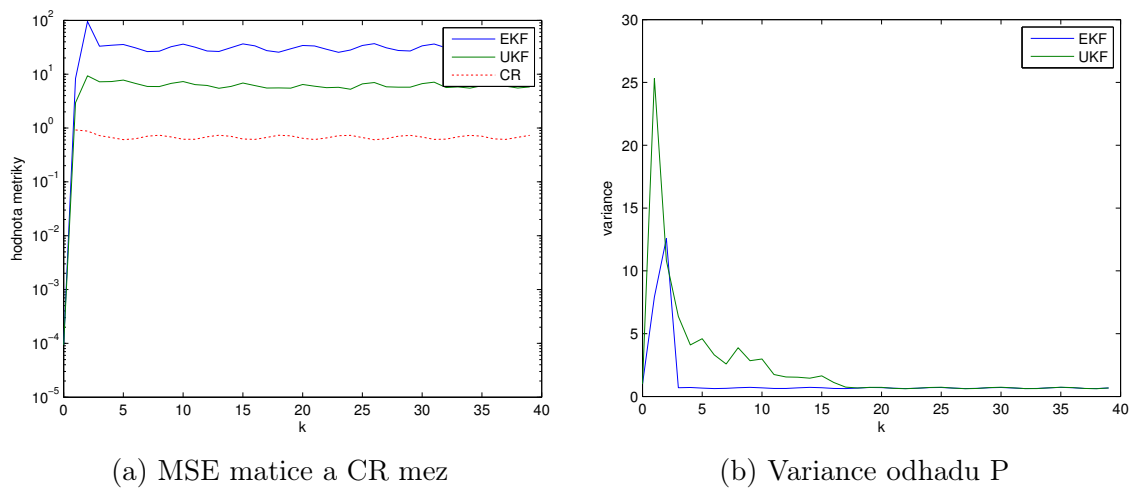
Obrázek 7.14: Průběh metrik BEEQ a EMER v čase.

O něco více vypovídající jsou metriky věrohodnosti ANEES a NCI, jejichž průběhy v čase jsou zachyceny na obrázku 7.15. Z obou obrázků je zřejmé, že UKF je o řád věrohodnější. To je způsobeno hlavně rozdílem mezi MSE maticemi obou estimátorů,

jejichž průběhy jsou zobrazeny na obrázku 7.16a. Dokonce ani není třeba mluvit o maticích (či dokonce použít skalární maticová hodnocení), protože dimenze stavu je 1, a tak se jedná o skalární hodnoty. Tamtéž je zobrazena CR mez, které se estimátory ani nepřiblíží. Posledním obrázkem je vývoj kovarianční matice odhadé dodané estimátory, který je předmětem obrázku 7.16b, kde opět platí, že se nejedná o matici, ale o skalární varianci chyby odhadu. Z obrázku je patrné, že si UKF na rozdíl od EKF alespoň několik kroků drží vyšší varianci chyby odhadu, což může být rozhodující faktor, proč je UKF lepším estimátorem.



Obrázek 7.15: Průběh metrik věrohodnosti v čase.



Obrázek 7.16: Průběh variance odhadu, MSE a CR meze v čase.

Kapitola 8

Závěr

Cílem této práce bylo představit a porovnat stávající metody hodnocení kvality odhadu stavu stochastických dynamických systémů diskrétních v čase. Práce se nejprve věnovala úloze odhadu stavu dynamických stochastických systémů a jejímu řešení globálními a lokálními estimátory. Lokální estimátory byly zkoumány podrobněji se zaměřením na algoritmy rozšířeného Kalmanova filtru a unscenovaného Kalmanova filtru. Dále byla pozornost soustředěna na metriky bodových odhadů poskytovaných lokálními nelineárními estimátory.

Metriky byly rozděleny na metriky hodnotící pouze chybu odhadu stavu a na metriky hodnotící i druhé momenty stavu či chyby odhadu stavu (např. kovarianční matici chyby odhadu dodanou estimátorem). Metriky hodnotící pouze chybu odhadu byly dále rozčleněny na metriky absolutní chyby odhadu, relativní chyby odhadu a na míry četnosti. Z každé této skupiny bylo uvedeno několik metrik, přičemž důraz byl vždy kladen na osvětlení jejich výhod, nevýhod a na jejich typické nasazení. Paralelně s tím byly metriky zkoumány z hlediska tolerování, resp. protežování malých či velkých chyb. Každá metrika tím byla klasifikována jako optimistická (tolerance velkých chyb), pesimistická (protežování velkých chyb) nebo vyvážená (ani optimistická ani pesimistická).

Druhou velkou skupinu tvořily metriky hodnotící i druhé momenty stavu či chyby odhadu stavu, tj. kovarianční matici chyby odhadu dodanou estimátorem a matici středních kvadratických chyb (MSE matici). Pozornost byla věnována nejen samostatnému ohodnocení těchto dvou matic, ale především jejich vzájemnému porovnání. Kovarianční matice chyby odhadu totiž představuje vlastní hodnocení chyby odhadu estimátorem a MSE matice je maticí reflektující skutečnou chybu odhadu. Porovnáním rozdílů mezi těmito dvěma maticemi tak byla hodnocena tzv. věrohodnost odhadu počtažmo estimátoru, tj. zda-li je estimátor optimistický (příliš sebevědomý), pesimistický (příliš nesmělý). Jako alternativní přístup k hodnocení kvality odhadu bylo představeno srovnání MSE matice s objektivním limitem poznatelnosti - Cramér-Raovou mezí.

Dále byla představena metodika hodnocení kvality odhadu stavu. V přehledných tabulkách byly shrnuty poznatky o všech metrikách ve smyslu jejich zařazení, výhod, nevýhod a jejich typické aplikace. Byl také ukázán postup při hodnocení odhadu několika metrikami zároveň. Aplikace těchto metrik byla předvedena na třech významných příkladech z různých oborů, aby byla dokladována univerzálnost a širě možné aplikace hodnocení kvality odhadu stavu. Metriky byly v těchto úlohách vystaveny typickým problémům, které je nutné řešit při hodnocení kvality, díky čemuž mohla být pro lepší

pochopení většina jejich vlastností ilustrována graficky.

Hodnocení kvality odhadu je přirozeným úkolem návrháře estimátoru, který ověřuje vlastnosti jím navrženého estimátoru ale i uživatele estimátoru, který se snaží aplikovat filtr na nějaký problém a ověřuje, zda jím vybraný estimátor je pro danou úlohu vhodný. Hodnocení kvality odhadu je nezbytné i při hodnocení výkonu vybrané sady estimátorů pro nějakou úlohu s cílem poskytnout doporučení pro výběr estimátoru. Avšak provést důsledné hodnocení kvality odhadu stavu pomocí jediné metriky je pouhou iluzí. Každá metrika se soustředí pouze na určité aspekty kvality odhadu, proto je důležité znát dostatečný počet metrik hodnotících různé aspekty. Zároveň by měl mít uživatel dostatečný přehled o této tématice. Splnit takovéto nároky není lehké, proto by stačilo mít k ruce alespoň souhrnnou práci na toto téma uvádějící tápajícího uživatele do tématu. Tato diplomová práce by mohla přispět k tomuto účelu. Už jen proto, že souhrnu podobného tématu není v české literatuře vůbec věnována pozornost, a v zahraniční literatuře zažívá toto téma po letech opomíjení nefalšovaný boom.

Literatura

- [1] Anderson, B.; Moore, J.: *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.
- [2] Arasaratnam, I.; Haykin, S.: Cubature Kalman Filters. *IEEE Transactions on Automatic Control*, ročník 54, č. 6, červen 2009: s. 1254–1269.
- [3] Arasaratnam, I.; Haykin, S.; Elliott, R.: Discrete-Time Nonlinear Filtering Algorithms Using Gauss-Hermite Quadrature. *Proceedings of the IEEE*, ročník 95, č. 5, květen 2007: s. 953–977.
- [4] Bar-Shalom, Y.; Birmiwal, K.: Consistency and Robustness of PDAF for Target Tracking in Cluttered Environments. *Automatica*, ročník 19, 1983: s. 431 – 437.
- [5] Bar-Shalom, Y.; Li, X. R.; Kirubarajan, T.: *Estimation with Applications to Tracking and Navigation: Theory, Algorithms, and Software*. New York: Wiley, 2001.
- [6] Blasch, E.; Rice, A.; Yang, C.; aj.: Relative Track Metrics to Determine Model Mismatch. In *Aerospace and Electronics Conference*, červenec 2008, s. 257–264.
- [7] Candy, J. V.: *Signal Processing: Model Based Approach*. New York, NY, USA: McGraw-Hill, Inc., 1986.
- [8] Doucet, A.; Freitas, N.; Gordon, N.: An Introduction to Sequential Monte Carlo Methods. In *Sequential Monte Carlo Methods in Practice*, Statistics for Engineering and Information Science, Springer New York, 2001, s. 3–14.
- [9] Duník, J.: *Návrh nelineárního systému odhadu v úlohách filtrace, predikce a vyhlazování*. Dizertační práce, ZČU v Plzni, 2007.
- [10] Duník, J.; Straka, O.; Šimandl, M.: Stochastic Integration Filter. *IEEE Transactions on Automatic Control*, ročník 58, č. 6, červen 2013: s. 1561–1566.
- [11] Haseltine, E. L.; Rawlings, J. B.: Critical Evaluation of Extended Kalman Filtering and Moving-Horizon Estimation. Technická zpráva, 2003.
- [12] Ito, K.; Xiong, K.: Gaussian Filters for Nonlinear Filtering Problems. *IEEE Transactions on Automatic Control*, ročník 45, 1999: s. 910–927.
- [13] Jazwinski, A. H.: *Stochastic Processes and Filtering Theory*. New York: Academic Press, 1970.

- [14] Julier, S. J.; Uhlmann, J. K.: Unscented Filtering and Nonlinear Estimation. *Proceedings of the IEEE*, ročník 92, č. 3, 2004: s. 401–422.
- [15] Kerr, T.: Status of CR-like Lower Bounds for Nonlinear Filtering. *IEEE Transactions on Aerospace and Electronic Systems*, ročník 25, č. 5, září 1989: s. 590–601.
- [16] Kitagawa, G.: Non-Gaussian State-Space Modeling of Nonstationary Time Series. *Journal of the American statistical association*, ročník 82, č. 400, 1987: s. 1032–1041.
- [17] Kolås, S.; Foss, B. A.; Schei, T. S.: Constrained Nonlinear State Estimation Based on the UKF Approach. *Computers & Chemical Engineering*, ročník 33, 2009: s. 1386–1401.
- [18] Kramer, S. C.; Sorenson, H. W.: Recursive Bayesian Estimation Using Piece-Wise Constant Approximations. *Automatica*, ročník 24, č. 6, 1988: s. 789 – 801.
- [19] Li, X. R.; Li, X.-B.: Common Fallacies in Applying Hypothesis Testing. In *11th International Conference on Information Fusion*, červen 2008, s. 1–8.
- [20] Li, X. R.; Zhao, Z.: Relative Error Measures for Evaluation of Estimation Algorithms. In *8th International Conference on Information Fusion*, ročník 1, červenec 2005.
- [21] Li, X. R.; Zhao, Z.: Evaluation of Estimation Algorithms Part I: Incomprehensive Measures of Performance. *IEEE Transactions on Aerospace and Electronic Systems*, ročník 42, č. 4, říjen 2006: s. 1340–1358.
- [22] Li, X. R.; Zhao, Z.: Measuring Estimator’s Credibility: Noncredibility Index. In *9th International Conference on Information Fusion*, červenec 2006, s. 1–8.
- [23] Li, X. R.; Zhao, Z.; Duan, Z.-S.: Error Spectrum and Desirability Level for Estimation Performance Evaluation. In *Workshop on Estimation, Tracking and Fusion: A Tribute to Fred Daum*, Moterey, CA, USA, květen 2007.
- [24] Li, X. R.; Zhao, Z.; Jilkov, V. P.: Practical Measures and Test for Credibility of an Estimator. In *Workshop on Estimation, Tracking, and Fusion - A Tribute to Yaakov Bar-Shalom*, Moterey, CA, USA, květen 2001, s. 481–495.
- [25] Li, X. R.; Zhao, Z.; Jilkov, V. P.: Estimator’s Credibility and Its Measures. In *Proc. IFAC 15th World Congress*, 2002.
- [26] Li, X. R.; Zhao, Z.; Li, X.-B.: Evaluation of Estimation Algorithms: Credibility Tests. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, ročník 42, č. 1, leden 2012: s. 147–163.
- [27] Meyer, C. D. (editor): *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2000.
- [28] Nørgaard, M.; Poulsen, N. K.; Ravn, O.: New Developments in State Estimation for Nonlinear Systems. *Automatica*, ročník 36, č. 11, listopad 2000: s. 1627–1638.

- [29] Perea, L.; How, J.; Breger, L.; aj.: Nonlinearity in Sensor Fusion: Divergence Issues. In *Proceedings of AIAA Guidance, Navigation, and Control Conf*, 2007.
- [30] Pitman, E. J. G.: The Closest Estimates of Statistical Parameters. *Mathematical Proceedings of the Cambridge Philosophical Society*, ročník 33, duben 1937: s. 212–222.
- [31] Sarmavuori, J.; Sarkka, S.: Fourier-Hermite Kalman Filter. *IEEE Transactions on Automatic Control*, ročník 57, č. 6, červen 2012: s. 1511–1515.
- [32] Sorenson, H.; Alspach, D.: Recursive Bayesian Estimation Using Gaussian Sums. *Automatica*, ročník 7, č. 4, 1971: s. 465 – 479.
- [33] Straka, O.; Duník, J.; Šimandl, M.: Performance Evaluation of Local State Estimation Methods in Bearings-Only Tracking Problems. In *14th International Conference on Information Fusion*, červenec 2011, s. 1–8.
- [34] Šimandl, M.: *Identifikace systémů a filtrace*. Plzeň: ZČU Plzeň, 1995.
- [35] Šimandl, M.: *Lecture Notes on State Estimation of Nonlinear Non-Gaussian Stochastic Systems*. Plzeň: ZČU v Plzni, 2006.
- [36] Šimandl, M.: Odhad stavu stochastických systémů. Centrum pro rozvoj výzkumu pokročilých řídicích a senzorických technologií, prosinec 2010, brožura k semináři.
- [37] Šimandl, M.; Královec, J.; Tichavský, P.: Filtering, Predictive, and Smoothing Cramér–Rao Bounds for Discrete-Time Nonlinear Dynamic Systems. *Automatica*, ročník 37, č. 11, 2001: s. 1703 – 1716.
- [38] Van Trees, H. L.: *Detection, Estimation, and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise*. Melbourne, FL, USA: Krieger Publishing Co., Inc., 1992.
- [39] Varga, R.: *Matrix Iterative Analysis*. Springer Series in Computational Mathematics, Springer, 2009.
- [40] Welch, G.; Bishop, G.: An Introduction to the Kalman Filter. Technická zpráva, Chapel Hill, NC, USA, 1995.
- [41] Wu, Y.; Hu, D.; Wu, M.; aj.: Unscented Kalman Filtering for Additive Noise Case: Augmented versus Nonaugmented. *Signal Processing Letters, IEEE*, ročník 12, č. 5, květen 2005: s. 357–360.
- [42] Yang, C.; Kaplan, L.; Blasch, E.: Performance Measures of Covariance and Information Matrices in Resource Management for Target State Estimation. *IEEE Transactions on Aerospace and Electronic Systems*, ročník 48, č. 3, červenec 2012: s. 2594–2613.
- [43] Yin, H.; Lan, J.; Li, X. R.: New Robust Metrics of Central Tendency for Estimation Performance Evaluation. In *15th International Conference on Information Fusion*, červenec 2012, s. 2020–2027.

- [44] Yin, H.; Lan, J.; Li, X. R.: Measures for Ranking Estimation Performance Based on Single or Multiple Performance Metrics. In *16th International Conference on Information Fusion*, červenec 2013, s. 453–460.