

# Gaussian Process Models in System Identification and State Estimation

*Ing. Jakub Průher*

Posudek disertační práce

Disertační práce Ing. Jakuba Průhera se zabývá využitím gaussových procesů při identifikaci dynamických systémů a odhadu stavu. Práce pokrývá několik úzce souvisejících témat: použití gaussovských procesů pro rekurzivní identifikaci dynamických systémů, zpřesnění metod transformace momentů a využití informace o gradientu při numerické integraci. Sjednocujícím faktorem pro tato témata je využití gaussových procesů a bayesovského přístupu.

## *Zhodnocení významu disertační práce pro obor*

Zvolené téma práce je aktuální, a to jak z hlediska rozvoje teoretického aparátu, tak z hlediska aplikací. Využití gaussových procesů spolu s metodami Monte Carlo významně rozšiřuje praktickou aplikovatelnost bayesovského přístupu na nelineární a negaussovské problémy. Aktuálnost tématu a intenzita výzkumu v této oblasti je dokumentovaná také kvalitním seznamem referencí, obsahujícím více než 170 položek.

## *Vyjádření k postupu řešení problému, použitým metodám a splnění určeného cíle*

Postup řešení problému, použitý doktorandem, je logický a dobře strukturovaný. V kapitole 2 doktorand shrnuje současný stav poznání v oblasti identifikace dynamických systémů. V kapitole 3 uvádí sjednocující pohled na metody nelineární filtrace s využitím transformace momentů. Obě tyto kapitoly jsou formulovány velmi přehledně a čtivě.

V kapitole 3 doktorand formuluje tři specifické cíle práce, a to

1. Využití Gaussovských procesů pro rekurzivní identifikaci dynamických systémů
2. Zpřesnění metod transformace momentů při numerické integraci využívající Gaussovských procesů
3. Využití informace o gradientu při numerické integraci využívající Gaussovských procesů.

Jednotlivé cíle práce jsou detailně rozpracovány v kapitolách 5, 6, 7. Použité metody založené na využití gaussových procesů a bayesovské statistiky jsou vhodným nástrojem k dosažení deklarovaných cílů práce.

### ***Stanovisko k výsledkům disertační práce a k původnímu přínosu autora***

Kapitola 5 prezentuje původní výsledky autora v oblasti využití gaussových procesů pro rekurzivní identifikaci dynamických systémů. Výsledky jsou založeny na konceptu rekurzivní GPR, vycházející z práce Huber (2013). Formulace problému (5.1.1.) je dostatečně obecná a výsledná kovarianční funkce (5.2.7) umožňuje dobře zahrnout apriorní informaci o funkcích  $f$  a  $g$ . Tuto alternativu však autor nevyužívá. K popsání implementaci duálního řízení mám řadu připomínek, které jsou uvedeny dále.

Výsledky autora v oblasti transformace momentů při numerické integraci využívající Gaussovských procesů prezentované v kapitole 6 jsou podle mého názoru podstatně kvalitnější. V této části jsem nenalezl žádné nedostatky, naopak oceňuji zajímavá inovativní kritéria, která autor použil pro porovnání jednotlivých metod (např. indikátor INC). Výsledky testů jednoznačně demonstrují přínos autora, za zvlášť významné považuji dosažení realistických hodnot kovarianční matice odhadu ilustrované např. na obr. 6.2.6.

Výsledky prezentované v kapitole 7 jsou poměrně přímočarou aplikací lineárního operátoru na gaussov proces, která se promítá na funkci jádra. Původním přínosem autora je detailní analýza souvislosti navrženého postupu s linearizací v odst. 7.2.1. Příklad 7.3.1 výstižným způsobem ilustruje zpřesnění, které využití informace o gradientu může přinést.

### ***Vyjádření k systematické, přehlednosti, formální úpravě a jazykové úrovni práce***

Uspořádání práce je systematické a logické, výklad je srozumitelný, rozsah jednotlivých kapitol přiměřený.

Práce má dobrou jazykovou i grafickou úroveň a obsahuje pouze drobné formální nedostatky (zaměněný horní a dolní index v Definicí 1 na str. 31, prohozený popis odezev na obr. 5.4.1 a 3 na str. 57, inicializace parametrů hodnotou  $\theta = \log([0 \ 0 \ 0 \ 0])$  na str. 67).

### ***Vyjádření k publikacím autora***

Výsledky doktoranda byly publikovány v časopisech a na kvalitních mezinárodních konferencích. Významným úspěchem je publikace v prestižním časopise *IEEE Transactions on Automatic Control*. Oceňuji také, že několik publikací autora vzniklo v rámci mezinárodní spolupráce s Aalto University ve Finsku.

### ***Připomínky oponenta a náměty k odborné diskusi***

1. Jaká je hodnota parametru  $\nu$  ve vztahu (3.1.12b)? Vzhledem k tomu, že veličina  $\beta$  definovaná (3.1.12c) má rozdělení  $\chi^2$  s  $d_z$  stupni volnosti, je ve vztahu pro kovarianční matici (3.1.12b) střední hodnota multiplikativního faktoru  $E\{(\nu-2+\beta)/(\nu-2+d_z)\} = 1$ . Má uvažování Studentova rozdělení opravdu významný vliv na chování odhadu (i pro konečné hodnoty parametru  $\nu$ ) ?

2. Pokud matice  $\mathbf{P}$  ve vztahu (3.1.13) odpovídá druhému centrálnímu momentu, jak je aplikována metoda momentů při uvažované změně počtu stupňů volnosti  $v$ ?
3. Inicializace hyper parametrů  $\theta_0 = [0 \ 0 \ 0 \ 0]$  uvedená na str. 56 nedává dobrý smysl. Funkce  $k(\Phi_i, \Phi_j, \theta_0)$  nesplňuje požadavky na funkci jádra (kovarianční matice  $\mathbf{K}(\theta_0)$  je nulová) a gradient logaritmické věrohodnostní funkce (2.2.19) není v tomto případě dobře definován (obsahuje výraz  $\mathbf{K}(\theta_0)^{-1}$  )
4. Předpoklady uvedené v odstavci 5.1 nemusí být dostatečné. Je známo, že strategie *minimum variance control* (5.6.1) nevede v některých případech na stabilizující zákon řízení.
5. Autor využívá GPR k identifikaci časově invariantního modelu (5.1.1). Vhodná volba hyper parametrů při využití jader typu „squared exponential“ pro neparametrický odhad funkce  $y = f(x)$  by měla umožnit dobrou aproximaci v místech s největší „křivostí“  $\delta^2 f(\cdot)/\delta x^2$ . Tato vlastnost není časově proměnná, proto nepovažují průběžnou aktualizaci hyper parametrů maximalizující marginální věrohodnost (5.2.20) za dobře zdůvodněnou. Můžete porovnat odezvy modelů na obr. 5.4.2 pro pevné hodnoty hyper parametrů  $\theta$  (např. ustálené hodnoty z Obr. 5.4.3.) – má použití časově proměnných hyper parametrů nějaký přínos?
6. Odvození vztahů (2.2.24) ... (2.2.28), které umožňují aproximovat informaci z množiny dat rostoucí dimenze pomocí středních hodnot a rozptylů pro předdefinovanou množinu bázových vektorů  $\Phi_i$  pevné dimenze, vychází z fixních jádrových funkcí  $k(\Phi_i, \Phi_j)$ . Pokud autor používá při rekurzivní regresi proměnné hyper parametry, měl by detailně odvodit tyto vztahy a zapracovat do nich proměnné jádrové funkce  $k(\Phi_i, \Phi_j, \theta_k)$ , nebo explicitně dokázat, že uvedená aproximace není na změnách funkce jádra závislá. Jinak není navrhované použití rekurzivní GPR korektní.
7. Volba hyper parametrů na základě největší „křivosti“  $\delta^2 f(\cdot)/\delta x^2$  umožňuje také kvalifikovanou volbu množiny bázových vektorů  $\Phi_i$ . Pokud není při volbě této báze respektován „vzorkovací teorém“, může být volba hyper parametrů (5.3.9) na základě pozorování  $\mu_k$  zcela zavádějící.
8. Zdůvodnění kritéria (5.6.3) na str. 62 není správné. Pokud by cílem druhého kritéria bylo penalizovat chybu predikce, vedlo by to typicky ke snížení amplitudy řízení. Vzhledem k zápornému znaménku kritéria (5.6.3) je naopak chyba predikce maximalizována. V parametrickém případě lze druhé kritérium zdůvodnit např. maximálním nárůstem informační matice parametrů. Bylo by možné najít podobnou interpretaci pro neparametrické modely GPR?
9. Vztah (5.6.7) je chybný, platí  $E\{(r-y)^2\} = (r-E\{y\})^2 + \text{cov}\{y\}$
10. Odvození výsledného zákona řízení na straně 64 a 65 je nesrozumitelné. Pokud hledáme řízení dle (5.6.4) v množině  $\Omega_k = [u_k - \delta, u_k + \delta]$ , měla by výsledná hodnota řízení (5.6.14) mít tvar  $u_k - \delta \text{sign}(dJ^a/du_k)$ .

11. Při vyhodnocení kvality řízení v tabulce 5.6.1 by bylo zajímavé doplnit porovnání s *certainty equivalent* regulátorem.

12. Operátor integrace je stejně jako operátor derivování lineárním operátorem. Bylo by možné využít k výpočtu integrálu Gaussova procesu integrální transformaci jádra (podobně jako diferenciální transformaci ve vztazích (7.1.14) a (7.1.15)) ?

### *Celkové zhodnocení práce*

Práce Ing. Jakuba Prühera obsahuje originální výsledky autora, které jsou přínosem pro vědeckou komunitu v oboru a tedy splňuje nároky na udělení akademického titulu „doktor“ v oboru *Kybernetika*. Zároveň konstatuji, že jsem v práci našel závažné nedostatky, ale předpokládám, že doktorand při obhajobě bude schopen ukázat, jak tyto nedostatky odstranit.

Doporučuji disertační práci Ing. Ing. Jakuba Prühera k obhajobě.

V Praze 26. 7. 2018



prof. Ing. Vladimír Havlena, CSc.



Karlsruhe Institute of Technology

KIT-Campus South | Intelligent Sensor-Actuator-Systems - ISAS  
P.O.Box 6980 | 76049 Karlsruhe, Germany

Ondrej Straka  
Department of Cybernetics  
University of West Bohemia  
Univerzitni 8  
306 14 Plzen  
Czech Republic

Institute for Anthropomatics and Robotics  
Intelligent Sensor-Actuator-Systems Lab - ISAS

Head: Prof. Dr.-Ing. Uwe D. Hanebeck

Kaiserstraße 12  
76131 Karlsruhe, Germany

Phone: +49 721-608-4-3909  
Fax: +49 721-608-4-5340  
Email: uwe.hanebeck@kit.edu  
Web: <http://isas.uka.de>

Official in charge:  
Our reference:  
Date: 2018-12-17

Report on the Doctoral Thesis of  
**Ing. Jakob Prüher**  
entitled  
**Gaussian Process Models in System Identification and State Estimation**  
supervised by  
**Doc. Ing. Ondřej Straka, Ph.D.**

### Context of the Thesis

The thesis of Mr. Prüher is concerned with Gaussian process regression, which is a nonparametric Bayesian method for estimating nonlinear models based on noisy data. Originally developed under the name "kriging" in geostatistics in the 1960's, Gaussian processes (GP) are an established and popular method in the machine learning community.

Two applications of GPs are considered in this thesis:

- 1) The first application is the recursive identification of discrete-time nonlinear stochastic time-invariant dynamical systems from noisy measurements of system inputs and outputs.
- 2) The second application is the calculation of moments in local nonlinear filters.

In the context of these applications, three goals are formulated:

- 1) Employing recursive Gaussian process regression for the recursive identification of nonlinear dynamic systems,
- 2) improving current moment transformation methods in local nonlinear filters, and
- 3) using derivative information for reducing the uncertainty in Bayesian quadrature.



Karlsruhe Institute of Technology (KIT)  
Kaiserstr. 12  
76131 Karlsruhe, Germany  
US-IdNr. DE266749428

President: Prof. Dr.-Ing. Holger Hanselka  
Vice Presidents: Prof. Dr. Thomas Hirth,  
Prof. Dr. Oliver Kraft, Christine von Vangerow,  
Prof. Dr. Alexander Wanner

LBBW/BW Bank  
IBAN: DE44 6005 0101 7495 5001 49  
BIC/SWIFT: SOLADEST600

LBBW/BW Bank  
IBAN: DE18 6005 0101 7495 5012 96  
BIC /SWIF: SOLADEST600

There are significant challenges associated with these goals, which include the theoretical understanding of the involved concepts and coping with the computational complexity of recursive Gaussian process regression.

### Summary of the Thesis

The present thesis consists of an introduction (chapter 1), six main chapters (chapter 2 to chapter 7), a conclusion in chapter 8, three appendices, and a bibliography with more than 160 references.

The introduction in **chapter 1** gives a short overview of Gaussian processes and applications, system identification, and state estimation. Finally, the structure of the thesis is shown.

The remaining part of the thesis is divided into two parts. The first part covers the basics and consists of chapter 2 and chapter 3. The second part covers the contributions and consists of chapters 4 to 7.

**Chapter 2** is devoted to the basics of system identification. It describes popular parametric system models in section 2.1. In section 2.2 nonparametric Gaussian processes are introduced, which are generalized to Student-t processes in section 2.3.

**Chapter 3** introduces states estimation by means of local nonlinear filtering, see section 3.1. The focus is on the transformation of moments, see section 3.2, which is identified as the central problem of local nonlinear filtering. Several standard moment transformation methods are shown. Bayesian quadrature is introduced in section 3.3.

**Chapter 4** discusses the two applications, system identification and moment calculation, and gives the three main goals of the thesis.

In **chapter 5**, the problem of recursive system identification with GP models is considered. The problem setup is given in section 5.1. The full GP model is presented in section 5.2, while section 5.3 introduces the recursive GP model. Numerical experiments are shown in section 5.4 and conclusions are drawn in section 5.5. The application of recursive GP models to functional dual adaptive control is presented in section 5.6.

**Chapter 6** is concerned with using Bayesian quadrature to calculate moment transforms. Section 6.1 introduces the general setup of Bayesian quadrature in the context of moment transforms. Section 6.2 focuses on Gaussian process quadrature moment transforms, while section 6.3. focuses on Student-t process quadrature moment transforms.

In **chapter 7**, the concept of Bayesian quadrature is generalized to using derivative information in the case that the integrand is differentiable. Section 7.1 discusses derivative information, while section 7.2 applies these quantities to Gaussian process quadrature moment transforms. The performance gain is demonstrated in section 7.2 by means of numerical evaluations. Section 7.4 concludes the chapter.

In **chapter 8**, the thesis is summarized and the main contributions are discussed. Section 8.1 then discusses challenges and gives an outlook to possible future work.

### Contributions of the Thesis

Mr. Prüher successfully applied Gaussian process regression methods to the recursive identification on nonlinear dynamical systems and to moment transformation in local nonlinear filters. In this context, he reached his three initial goals that include 1) employing recursive Gaussian process regression for the recursive identification of nonlinear dynamic systems, 2) improving current moment transformation methods in local nonlinear filters, and 3) using derivative information for reducing the uncertainty in Bayesian quadrature.

Goal 1) was solved in chapter 5 by using the recursive Gaussian process algorithm for system identification while proposing a faster kernel parameter optimization method for further reducing the computational complexity. Evaluations demonstrated a comparable performance to the full Gaussian process at a significantly lower computational cost, which was further underlined by using the proposed identification method in an adaptive control loop.

Goal 2) was attacked in chapter 6 by including the integral uncertainty of both Gaussian and Student-t process models into the moment transform. It was demonstrated that this proposed transform mechanism significantly improved the desired covariance estimates. The method is currently limited to low-dimensional system. In addition, it has to be mentioned that the performance of local nonlinear filters is generally limited by the assumption that state and measurement are jointly Gaussian distributed. Hence, even the optimum values for the moments would not result in the same performance as a Gaussian-assumed density filter (that does not assume joint Gaussianity).

Goal 3) was approached in chapter 7. It was found that using derivative information significantly reduces the integration errors in the Gaussian process quadrature moment transform.

### Evaluation


Mr. Prüher mastered the various challenges in his pursuit of solving his three main goals in the context of system identification and moment transforms. He showed a high level of competence and knowledge of the fields of system theory, state estimation, and machine learning. His theoretical concepts are well thought-out and have been converted to practical implementations. The methods have been thoroughly tested and evaluated in realistic simulations. The present thesis is a major step in the direction of the practical applicability of nonlinear estimation and control.

The thesis is clearly structured and very well formulated. The proposed methods are correctly derived, well documented, and easily comprehensible. Mr. Prüher published his results in 8 papers (7 as first author, 1 as second author).

With the present thesis, Ing. Jakub Prüher clearly demonstrated his abilities for independent scientific work and for the systematic presentation of his theoretical and practical results. In addition, Mr. Prüher provided new and original scientific contributions.

I strongly suggest to the Department of Cybernetics of the Faculty of Applied Science, University of West Bohemia the *continuation of the promotion procedure and recommend this thesis for the defense.*

Best regards,

A handwritten signature in black ink, consisting of several loops and a long horizontal stroke at the end.

U. D. Hanebeck  
(Univ.-Professor)