**Bayesian Filtering and Smoothing**

Filtering and smoothing methods are used to produce an accurate estimate of the state of a time-varying system based on multiple observational inputs (data). Interest in these methods has exploded in recent years, with numerous applications emerging in fields such as navigation, aerospace engineering, telecommunications, and medicine.

This compact, informal introduction for graduate students and advanced undergraduates presents the current state-of-the-art filtering and smoothing methods in a unified Bayesian framework. Readers learn what non-linear Kalman filters and particle filters are, how they are related, and their relative advantages and disadvantages. They also discover how state-of-the-art Bayesian parameter estimation methods can be combined with state-of-the-art filtering and smoothing algorithms.

The book's practical and algorithmic approach assumes only modest mathematical prerequisites. Examples include MATLAB computations, and the numerous end-of-chapter exercises include computational assignments. MATLAB/GNU Octave source code is available for download at www.cambridge.org/sarkka, promoting hands-on work with the methods.

SIMO SÄRKKÄ worked, from 2000 to 2010, with Nokia Ltd., Indagon Ltd., and the Nalco Company in various industrial research projects related to telecommunications, positioning systems, and industrial process control. Currently, he is a Senior Researcher with the Department of Biomedical Engineering and Computational Science at Aalto University, Finland, and Adjunct Professor with Tampere University of Technology and Lappeenranta University of Technology. In 2011 he was a visiting scholar with the Signal Processing and Communications Laboratory of the Department of Engineering at the University of Cambridge. His research interests are in state and parameter estimation in stochastic dynamic systems and, in particular, Bayesian methods in signal processing, machine learning, and inverse problems with applications to brain imaging, positioning systems, computer vision, and audio signal processing. He is a Senior Member of the IEEE.

INSTITUTE OF MATHEMATICAL STATISTICS TEXTBOOKS



IMS Textbooks give introductory accounts of topics of current concern suitable for advanced courses at master's level, for doctoral students and for individual study. They are typically shorter than a fully developed textbook, often arising from material created for a topical course. Lengths of 100–290 pages are envisaged. The books typically contain exercises.

# Bayesian Filtering and Smoothing

SIMO SÄRKKÄ
*Aalto University, Finland*

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107030657



First published 2013
3rd printing 2014

Printed in the United Kingdom by Clays, St Ives plc.

*A catalogue record for this publication is available from the British Library*

*Library of Congress Cataloguing in Publication data*

ISBN 978-1-107-03065-7 Hardback
ISBN 978-1-107-61928-9 Paperback

Additional resources for this publication at www.cambridge.org/sarkka

# Contents

# Preface

The aim of this book is to give a concise introduction to non-linear Kalman filtering and smoothing, particle filtering and smoothing, and to the related parameter estimation methods. Although the book is intended to be an introduction, the mathematical ideas behind all the methods are carefully explained, and a mathematically inclined reader can get quite a deep understanding of the methods by reading the book. The book is purposely kept short for quick reading.

The book is mainly intended for advanced undergraduate and graduate students in applied mathematics and computer science. However, the book is suitable also for researchers and practitioners (engineers) who need a concise introduction to the topic on a level that enables them to implement or use the methods. The assumed background is linear algebra, vector calculus, Bayesian inference, and MATLAB® programming skills.

As implied by the title, the mathematical treatment of the models and algorithms in this book is Bayesian, which means that all the results are treated as being approximations to certain probability distributions or their parameters. Probability distributions are used both to represent uncertainties in the models and for modeling the physical randomness. The theories of non-linear filtering, smoothing, and parameter estimation are formulated in terms of Bayesian inference, and both the classical and recent algorithms are derived using the same Bayesian notation and formalism. This Bayesian approach to the topic is far from new. It was pioneered by Stratonovich in the 1950s and 1960s – even before Kalman's seminal article in 1960. Thus the theory of non-linear filtering has been Bayesian from the beginning (see Jazwinski, 1970).

Chapter 1 is a general introduction to the idea and applications of Bayesian filtering and smoothing. The purpose of Chapter 2 is to briefly review the basic concepts of Bayesian inference as well as the basic numerical methods used in Bayesian computations. Chapter 3 starts with a step-by-step introduction to recursive Bayesian estimation via solving a linear regression problem in a recursive manner. The transition to Bayesian filtering and smoothing theory is explained by extending and generalizing the problem. The first Kalman filter of the book is also encountered in this chapter.

The Bayesian filtering theory starts in Chapter 4 where we derive the general Bayesian filtering equations and, as their special case, the celebrated Kalman filter. Non-linear extensions of the Kalman filter, the extended Kalman filter (EKF), the statistically linearized filter (SLF), and the unscented Kalman filter (UKF) are presented in Chapter 5. Chapter 6 generalizes these filters into the framework of Gaussian filtering. The Gauss–Hermite Kalman filter (GHKF) and cubature Kalman filter (CKF) are then derived from the general framework. Sequential Monte Carlo (SMC) based particle filters (PF) are explained in Chapter 7 by starting from the basic SIR filter and ending with Rao–Blackwellized particle filters (RBPF).

Chapter 8 starts with a derivation of the general (fixed-interval) Bayesian smoothing equations and then continues to a derivation of the Rauch–Tung–Striebel (RTS) smoother as their special case. In that chapter we also briefly discuss two-filter smoothing. The extended RTS smoother (ERTSS), statistically linearized RTS smoother (SLRTSS), and the unscented RTS smoother (URTSS) are presented in Chapter 9. The general Gaussian smoothing framework is presented in Chapter 10, and the Gauss–Hermite RTS smoother (GHRTSS) and the cubature RTS smoother (CRTSS) are derived as its special cases. We also discuss Gaussian fixed-point and fixed-lag smoothing in the same chapter. In Chapter 11 we start by showing how the basic SIR particle filter can be used to approximate the smoothing solutions with a small modification. We then introduce the numerically better backward-simulation particle smoother and the reweighting (or marginal) particle smoother. Finally, we discuss the implementation of Rao–Blackwellized particle smoothers.

Chapter 12 is an introduction to parameter estimation in state space models concentrating on optimization and expectation–maximization (EM) based computation of maximum likelihood (ML) and maximum a posteriori (MAP) estimates, as well as to Markov chain Monte Carlo (MCMC) methods. We start by presenting the general methods and then show how Kalman filters and RTS smoothers, non-linear Gaussian filters and RTS smoothers, and finally particle filters and smoothers, can be used to compute or approximate the quantities needed in implementation of parameter estimation methods. This leads to, for example, classical EM algorithms for state space models, as well as to particle EM and particle MCMC methods. We also discuss how Rao–Blackwellization can sometimes be used to help parameter estimation.

Chapter 13 is an epilogue where we give some general advice on the selection of different methods for different purposes. We also discuss and give references to various technical points and related topics that are important, but did not fit into this book.

Each of the chapters ends with a range of exercises, which give the reader hands-on experience in implementing the methods and in selecting the appropriate method for a given purpose. The MATLAB® source code needed in the exercises as well as various other material can be found on the book's web page at `www.cambridge.org/sarkka`.

This book is an outgrowth of lecture notes of courses that I gave during the years 2009–2012 at Helsinki University of Technology, Aalto University, and Tampere University of Technology, Finland. Most of the text was written while I was working at the Department of Biomedical Engineering and Computational Science (BECS) of Aalto University (formerly Helsinki University of Technology), but some of the text was written during my visit to the Department of Engineering at the University of Cambridge, UK. I am grateful to the former Centre of Excellence in Computational Complex Systems Research of the Academy of Finland, BECS, and Aalto University School of Science for providing me with the research funding which made this book possible.

I would like to thank Jouko Lampinen and Aki Vehtari from BECS for giving me the opportunity to do the research and for co-operation which led to this book. Arno Solin, Robert Piché, Juha Sarmavuori, Thomas Schön, Pete Bunch, and Isambi S. Mbalawata deserve thanks for careful checking of the book and for giving a lot of useful suggestions for improving the text. I am also grateful to Jouni Hartikainen, Ville Väänänen, Heikki Haario, and Simon Godsill for research co-operation that led to improvement of my understanding of the topic as well as to the development of some of the methods which now are explained in this book. I would also like to thank Diana Gillooly from Cambridge University Press and series editor Susan Holmes for suggesting the publication of my lecture notes in book form. Finally, I am grateful to my wife Susanne for her support and patience during the writing of this book.

*Simo Särkkä*
Vantaa, Finland

# Symbols and abbreviations

## General notation

| | |
|---|---|
| $a, b, c, x, t, \alpha, \beta$ | Scalars |
| $\mathbf{a}, \mathbf{f}, \mathbf{s}, \mathbf{x}, \mathbf{y}, \boldsymbol{\alpha}, \boldsymbol{\beta}$ | Vectors |
| $\mathbf{A}, \mathbf{F}, \mathbf{S}, \mathbf{X}, \mathbf{Y}$ | Matrices |
| $\mathcal{A}, \mathcal{F}, \mathcal{S}, \mathcal{X}, \mathcal{Y}$ | Sets |
| $\mathbb{A}, \mathbb{F}, \mathbb{S}, \mathbb{X}, \mathbb{Y}$ | Spaces |

## Notational conventions

| | |
|---|---|
| $\mathbf{A}^\mathsf{T}$ | Transpose of matrix |
| $\mathbf{A}^{-1}$ | Inverse of matrix |
| $\mathbf{A}^{-\mathsf{T}}$ | Inverse of transpose of matrix |
| $[\mathbf{A}]_i$ | $i$th column of matrix $\mathbf{A}$ |
| $[\mathbf{A}]_{ij}$ | Element at $i$th row and $j$th column of matrix $\mathbf{A}$ |
| $\lvert a \rvert$ | Absolute value of scalar $a$ |
| $\lvert \mathbf{A} \rvert$ | Determinant of matrix $\mathbf{A}$ |
| $\mathrm{d}\mathbf{x}/\mathrm{d}t$ | Time derivative of $\mathbf{x}(t)$ |
| $\frac{\partial g_i(\mathbf{x})}{\partial x_j}$ | Partial derivative of $g_i$ with respect to $x_j$ |
| $(a_1, \ldots, a_n)$ | Column vector with elements $a_1, \ldots, a_n$ |
| $(a_1 \ \cdots \ a_n)$ | Row vector with elements $a_1, \ldots, a_n$ |
| $(a_1 \ \cdots \ a_n)^\mathsf{T}$ | Column vector with elements $a_1, \ldots, a_n$ |
| $\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}}$ | Gradient (column vector) of scalar function $g$ |
| $\frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}$ | Jacobian matrix of vector valued function $\mathbf{x} \to \mathbf{g}(\mathbf{x})$ |
| $\mathrm{Cov}[\mathbf{x}]$ | Covariance $\mathrm{Cov}[\mathbf{x}] = \mathrm{E}[(\mathbf{x} - \mathrm{E}[\mathbf{x}])\,(\mathbf{x} - \mathrm{E}[\mathbf{x}])^\mathsf{T}]$ of the random variable $\mathbf{x}$ |
| $\mathrm{diag}(a_1, \ldots, a_n)$ | Diagonal matrix with diagonal values $a_1, \ldots, a_n$ |
| $\sqrt{\mathbf{P}}$ | Matrix such that $\mathbf{P} = \sqrt{\mathbf{P}}\,\sqrt{\mathbf{P}}^\mathsf{T}$ |
| $\mathrm{E}[\mathbf{x}]$ | Expectation of $\mathbf{x}$ |
| $\mathrm{E}[\mathbf{x} \mid \mathbf{y}]$ | Conditional expectation of $\mathbf{x}$ given $\mathbf{y}$ |

| | |
|---|---|
| $\int f(\mathbf{x})\, \mathrm{d}\mathbf{x}$ | Lebesgue integral of $f(\mathbf{x})$ over the space $\mathbb{R}^n$ |
| $p(\mathbf{x})$ | Probability density of continuous random variable $\mathbf{x}$ or probability of discrete random variable $\mathbf{x}$ |
| $p(\mathbf{x} \mid \mathbf{y})$ | Conditional probability density or conditional probability of $\mathbf{x}$ given $\mathbf{y}$ |
| $p(\mathbf{x}) \propto q(\mathbf{x})$ | $p(\mathbf{x})$ is proportional to $q(\mathbf{x})$, that is, there exists a constant $c$ such that $p(\mathbf{x}) = c\, q(\mathbf{x})$ for all values of $\mathbf{x}$ |
| $\mathrm{tr}\, \mathbf{A}$ | Trace of matrix $\mathbf{A}$ |
| $\mathrm{Var}[x]$ | Variance $\mathrm{Var}[x] = \mathrm{E}[(x - \mathrm{E}[x])^2]$ of the scalar random variable $x$ |
| $x \gg y$ | $x$ is much greater than $y$ |
| $x_{i,k}$ | $i$th component of vector $\mathbf{x}_k$ |
| $\mathbf{x} \sim p(\mathbf{x})$ | Random variable $\mathbf{x}$ has the probability density or probability distribution $p(\mathbf{x})$ |
| $\mathbf{x} \triangleq \mathbf{y}$ | $\mathbf{x}$ is defined to be equal to $\mathbf{y}$ |
| $\mathbf{x} \approx \mathbf{y}$ | $\mathbf{x}$ is approximately equal to $\mathbf{y}$ |
| $\mathbf{x} \simeq \mathbf{y}$ | $\mathbf{x}$ is assumed to be approximately equal to $\mathbf{y}$ |
| $\mathbf{x}_{0:k}$ | Set or sequence containing the vectors $\{\mathbf{x}_0, \ldots, \mathbf{x}_k\}$ |
| $\dot{\mathbf{x}}$ | Time derivative of $\mathbf{x}(t)$ |

## Symbols

| | |
|---|---|
| $\alpha$ | Parameter of the unscented transform or pendulum angle |
| $\alpha_i$ | Acceptance probability in an MCMC method |
| $\bar{\alpha}_*$ | Target acceptance rate in an adaptive MCMC |
| $\beta$ | Parameter of the unscented transform |
| $\delta(\cdot)$ | Dirac delta function |
| $\delta\mathbf{x}$ | Difference of $\mathbf{x}$ from the mean $\delta\mathbf{x} = \mathbf{x} - \mathbf{m}$ |
| $\Delta t$ | Sampling period |
| $\Delta t_k$ | Length of the time interval $\Delta t_k = t_{k+1} - t_k$ |
| $\varepsilon_k$ | Measurement error at the time step $k$ |
| $\boldsymbol{\varepsilon}_k$ | Vector of measurement errors at the time step $k$ |
| $\boldsymbol{\theta}$ | Vector of parameters |
| $\boldsymbol{\theta}_k$ | Vector of parameters at the time step $k$ |
| $\boldsymbol{\theta}^{(n)}$ | Vector of parameters at iteration $n$ of the EM-algorithm |
| $\boldsymbol{\theta}^{(i)}$ | Vector of parameters at iteration $i$ of the MCMC-algorithm |
| $\boldsymbol{\theta}^*$ | Candidate point in the MCMC-algorithm |
| $\hat{\boldsymbol{\theta}}^{\mathrm{MAP}}$ | Maximum a posteriori (MAP) estimate of parameter $\boldsymbol{\theta}$ |
| $\kappa$ | Parameter of the unscented transform |

| | |
|---|---|
| $\lambda$ | Parameter of the unscented transform |
| $\lambda'$ | Parameter of the unscented transform |
| $\lambda''$ | Parameter of the unscented transform |
| $\boldsymbol{\mu}_k$ | Predicted mean of measurement $\mathbf{y}_k$ in a Kalman/Gaussian filter at the time step $k$ |
| $\boldsymbol{\mu}_{\mathrm{L}}$ | Mean in the linear approximation of a non-linear transform |
| $\boldsymbol{\mu}_{\mathrm{M}}$ | Mean in the Gaussian moment matching approximation |
| $\boldsymbol{\mu}_{\mathrm{Q}}$ | Mean in the quadratic approximation |
| $\boldsymbol{\mu}_{\mathrm{S}}$ | Mean in the statistical linearization approximation |
| $\boldsymbol{\mu}_{\mathrm{U}}$ | Mean in the unscented approximation |
| $\pi(\cdot)$ | Importance distribution |
| $\sigma^2$ | Variance |
| $\sigma_i^2$ | Variance of noise component $i$ |
| $\boldsymbol{\Sigma}$ | Auxiliary matrix needed in the EM-algorithm |
| $\boldsymbol{\Sigma}_i$ | Proposal distribution covariance in the Metropolis algorithm |
| $\varphi_k(\boldsymbol{\theta})$ | Energy function at the time step $k$ |
| $\Phi(\cdot)$ | A function returning the lower triangular part of its argument |
| $\boldsymbol{\Phi}$ | An auxiliary matrix needed in the EM-algorithm |
| $\xi$ | Unit Gaussian random variable |
| $\xi^{(i)}$ | $i$th scalar unit sigma point |
| $\boldsymbol{\xi}$ | Vector of unit Gaussian random variables |
| $\boldsymbol{\xi}^{(i)}$ | $i$th unit sigma point vector |
| $\boldsymbol{\xi}^{(i_1,\ldots,i_n)}$ | Unit sigma point in the multivariate Gauss–Hermite cubature |
| $\mathbf{a}$ | Action in decision theory, or a part of a mean vector |
| $\mathbf{a}_{\mathrm{o}}$ | Optimal action |
| $\mathbf{a}(t)$ | Acceleration |
| $\mathbf{A}$ | Dynamic model matrix in a linear time-invariant model, the lower triangular Cholesky factor of a covariance matrix, the upper left block of a covariance matrix, a coefficient in statistical linearization, or an arbitrary matrix |
| $\mathbf{A}_k$ | Dynamic model matrix (i.e., transition matrix) of the jump from step $k$ to step $k+1$ |
| $\mathbf{b}$ | The lower part of a mean vector, the offset term in statistical linearization, or an arbitrary vector |
| $\mathbf{B}$ | Lower right block of a covariance matrix, an auxiliary matrix needed in the EM-algorithm, or an arbitrary matrix |
| $\mathbf{B}_{j\mid k}$ | Gain matrix in a fixed-point or fixed-lag Gaussian smoother |
| $c$ | Scalar constant |
| $C(\cdot)$ | Cost or loss function |

| | |
|---|---|
| $\mathbf{C}$ | The upper right block of a covariance matrix, an auxiliary matrix needed in the EM-algorithm, or an arbitrary matrix |
| $\mathbf{C}_k$ | Cross-covariance matrix in a non-linear Kalman filter |
| $\mathbf{C}_\mathrm{L}$ | Cross-covariance in the linear approximation of a non-linear transform |
| $\mathbf{C}_\mathrm{M}$ | Cross-covariance in the Gaussian moment matching approximation of a non-linear transform |
| $\mathbf{C}_\mathrm{Q}$ | Cross-covariance in the quadratic approximation |
| $\mathbf{C}_\mathrm{S}$ | Cross-covariance in the statistical linearization approximation |
| $\mathbf{C}_\mathrm{U}$ | Cross-covariance in the unscented approximation |
| $d$ | Positive integer, usually dimensionality of the parameters |
| $d_i$ | Order of a monomial |
| $\mathrm{d}t$ | Differential of time variable $t$ |
| $\mathrm{d}\mathbf{x}$ | Differential of vector $\mathbf{x}$ |
| $\mathbf{D}$ | Derivative of the Cholesky factor, an auxiliary matrix needed in the EM-algorithm, or an arbitrary matrix |
| $\mathbf{D}_k$ | Cross-covariance matrix in a non-linear RTS smoother or an auxiliary matrix used in derivations |
| $\mathbf{e}_i$ | Unit vector in the direction of the coordinate axis $i$ |
| $\mathbf{f}(\cdot)$ | Dynamic transition function in a state space model |
| $\mathbf{F}_\mathbf{x}(\cdot)$ | Jacobian matrix of the function $\mathbf{x} \to \mathbf{f}(\mathbf{x})$ |
| $\mathbf{F}$ | Feedback matrix of a continuous-time linear state space model |
| $\mathbf{F}^{(i)}_{\mathbf{xx}}(\cdot)$ | Hessian matrix of $\mathbf{x} \to f_i(\mathbf{x})$ |
| $F[\cdot]$ | An auxiliary functional needed in the derivation of the EM-algorithm |
| $g$ | Gravitation acceleration |
| $g(\cdot)$ | An arbitrary function |
| $g_i(\cdot)$ | An arbitrary function |
| $\mathbf{g}(\cdot)$ | An arbitrary function |
| $\mathbf{g}^{-1}(\cdot)$ | Inverse function of $\mathbf{g}(\cdot)$ |
| $\tilde{\mathbf{g}}(\cdot)$ | Augmented function with elements $(\mathbf{x}, \mathbf{g}(\cdot))$ |
| $\mathbf{G}_k$ | Gain matrix in an RTS smoother |
| $\mathbf{G}_\mathbf{x}(\cdot)$ | Jacobian matrix of the function $\mathbf{x} \to \mathbf{g}(\mathbf{x})$ |
| $\mathbf{G}^{(i)}_{\mathbf{xx}}(\cdot)$ | Hessian matrix of $\mathbf{x} \to g_i(\mathbf{x})$ |
| $H_p(\cdot)$ | $p$th order Hermite polynomial |
| $\mathbf{H}$ | Measurement model matrix in a linear Gaussian model, or a Hessian matrix |
| $\mathbf{H}_k$ | Measurement model matrix at the time step $k$ in a linear Gaussian model |

| | |
|---|---|
| $\mathbf{H}_{\mathbf{x}}(\cdot)$ | Jacobian matrix of the function $\mathbf{x} \to \mathbf{h}(\mathbf{x})$ |
| $\mathbf{H}^{(i)}_{\mathbf{xx}}(\cdot)$ | Hessian matrix of $\mathbf{x} \to h_i(\mathbf{x})$ |
| $\mathbf{h}(\cdot)$ | Measurement model function in a state space model |
| $i$ | Integer valued index variable |
| $\mathbf{I}$ | Identity matrix |
| $I_i(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)})$ | An integral term needed in the EM-algorithm |
| $\mathbf{J}(\cdot)$ | Jacobian matrix |
| $k$ | Time step number |
| $\mathbf{K}_k$ | Gain matrix of a Kalman/Gaussian filter |
| $\mathbf{L}$ | Noise coefficient (i.e., dispersion) matrix of a continuous-time linear state space model |
| $\mathcal{L}(\cdot)$ | Likelihood function |
| $m$ | Dimensionality of a measurement, mean of the univariate Gaussian distribution, or the mass |
| $\mathbf{m}$ | Mean of a Gaussian distribution |
| $\tilde{\mathbf{m}}$ | Mean of an augmented random variable |
| $\mathbf{m}_k$ | Mean of a Kalman/Gaussian filter at the time step $k$ |
| $\mathbf{m}^{(i)}_k$ | Mean of the Kalman filter in the particle $i$ of RBPF at the time step $k$ |
| $\mathbf{m}^{(i)}_{0:T}$ | History of means of the Kalman filter in the particle $i$ of RBPF |
| $\tilde{\mathbf{m}}_k$ | Augmented mean at the time step $k$ or an auxiliary variable used in derivations |
| $\mathbf{m}^-_k$ | Predicted mean of a Kalman/Gaussian filter at the time step $k$ just before the measurement $\mathbf{y}_k$ |
| $\mathbf{m}^{-(i)}_k$ | Predicted mean of the Kalman filter in the particle $i$ of RBPF at the time step $k$ |
| $\tilde{\mathbf{m}}^-_k$ | Augmented predicted mean at the time step $k$ |
| $\mathbf{m}^{\mathrm{s}}_k$ | Mean computed by a Gaussian fixed-interval (RTS) smoother for the time step $k$ |
| $\mathbf{m}^{\mathrm{s},(i)}_{0:T}$ | History of means of the RTS smoother in the particle $i$ of RBPS |
| $\mathbf{m}_{k\mid n}$ | Conditional mean of $\mathbf{x}_k$ given $\mathbf{y}_{1:n}$ |
| $n$ | Positive integer, usually the dimensionality of the state |
| $n'$ | Augmented state dimensionality in a non-linear transform |
| $n''$ | Augmented state dimensionality in a non-linear transform |
| $N$ | Positive integer, usually the number of Monte Carlo samples |
| $\mathrm{N}(\cdot)$ | Gaussian distribution (i.e., normal distribution) |

| | |
|---|---|
| $p$ | Order of a Hermite polynomial |
| $P$ | Variance of the univariate Gaussian distribution |
| $\mathbf{P}$ | Covariance of the Gaussian distribution |
| $\tilde{\mathbf{P}}$ | Covariance of an augmented random variable |
| $\mathbf{P}_k$ | Covariance of a Kalman/Gaussian filter at the time step $k$ |
| $\mathbf{P}_k^{(i)}$ | Covariance of the Kalman filter in the particle $i$ of RBPF at the time step $k$ |
| $\mathbf{P}_{0:T}^{(i)}$ | History of covariances of the Kalman filter in the particle $i$ of RBPF |
| $\tilde{\mathbf{P}}_k$ | Augmented covariance at the time step $k$ or an auxiliary variable used in derivations |
| $\mathbf{P}_k^-$ | Predicted covariance of a Kalman/Gaussian filter at the time step $k$ just before the measurement $\mathbf{y}_k$ |
| $\tilde{\mathbf{P}}_k^-$ | Augmented predicted covariance at the time step $k$ |
| $\mathbf{P}_k^{-(i)}$ | Predicted covariance of the Kalman filter in the particle $i$ of RBPF at the time step $k$ |
| $\mathbf{P}_k^{\mathrm{s}}$ | Covariance computed by a Gaussian fixed-interval (RTS) smoother for the time step $k$ |
| $\mathbf{P}_{0:T}^{\mathrm{s},(i)}$ | History of covariances of the RTS smoother in the particle $i$ of RBPS |
| $\mathbf{P}_{k\mid n}$ | Conditional covariance of $\mathbf{x}_k$ given $\mathbf{y}_{1:n}$ |
| $q^{\mathrm{c}}$ | Spectral density of a white noise process |
| $q_i^{\mathrm{c}}$ | Spectral density of component $i$ of a white noise process |
| $q(\cdot)$ | Proposal distribution in the MCMC algorithm, or an arbitrary distribution in the derivation of the EM-algorithm |
| $q^{(n)}$ | Distribution approximation on the $n$th step of the EM-algorithm |
| $\mathbf{q}$ | Gaussian random vector |
| $\mathbf{q}_k$ | Gaussian process noise |
| $Q$ | Variance of scalar process noise |
| $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(n)})$ | An auxiliary function needed in the EM-algorithm |
| $\mathbf{Q}$ | Covariance of the process noise in a time-invariant model |
| $\mathbf{Q}_k$ | Covariance of the process noise at the jump from step $k$ to $k+1$ |
| $r_k$ | Scalar Gaussian measurement noise |
| $\mathbf{r}_k$ | Vector of Gaussian measurement noises |
| $R$ | Variance of scalar measurement noise |
| $\mathbf{R}$ | Covariance matrix of the measurement in a time-invariant model |

| | |
|---|---|
| $\mathbf{R}_k$ | Covariance matrix of the measurement at the time step $k$ |
| $\mathbb{R}$ | Space of real numbers |
| $\mathbb{R}^n$ | $n$-dimensional space of real numbers |
| $\mathbb{R}^{n\times m}$ | Space of real $n \times m$ matrices |
| $S$ | Number of backward-simulation draws |
| $\mathbf{S}_k$ | Innovation covariance of a Kalman/Gaussian filter at step $k$ |
| $\mathbf{S}_\mathrm{L}$ | Covariance in the linear approximation of a non-linear transform |
| $\mathbf{S}_\mathrm{M}$ | Covariance in the Gaussian moment matching approximation of a non-linear transform |
| $\mathbf{S}_\mathrm{Q}$ | Covariance in the quadratic approximation of a non-linear transform |
| $\mathbf{S}_\mathrm{S}$ | Covariance in the statistical linearization approximation of a non-linear transform |
| $\mathbf{S}_\mathrm{U}$ | Covariance in the unscented approximation of a non-linear transform |
| $t$ | Time variable $t \in [0, \infty)$ |
| $t_k$ | Time of the step $k$ (usually time of the measurement $y_k$) |
| $T$ | Index of the last time step, the final time of a time interval |
| $\mathcal{T}_k$ | Sufficient statistics |
| $u$ | Uniform random variable |
| $\mathbf{u}_k$ | Latent (non-linear) variable in a Rao–Blackwellized particle filter or smoother |
| $\mathbf{u}_k^{(i)}$ | Latent variable value in particle $i$ |
| $\mathbf{u}_{0:k}^{(i)}$ | History of latent variable values in particle $i$ |
| $U(\cdot)$ | Utility function |
| $\mathrm{U}(\cdot)$ | Uniform distribution |
| $v_k^{(i)}$ | Unnormalized weight in an SIR particle filter based likelihood evaluation |
| $\mathbf{v}_k$ | Innovation vector of a Kalman/Gaussian filter at step $k$ |
| $w^{(i)}$ | Normalized weight of the particle $i$ in importance sampling |
| $\tilde{w}^{(i)}$ | Weight of the particle $i$ in importance sampling |
| $w^{*(i)}$ | Unnormalized weight of the particle $i$ in importance sampling |
| $w_k^{(i)}$ | Normalized weight of the particle $i$ on step $k$ of a particle filter |
| $w_{k\mid n}^{(i)}$ | Normalized weight of a particle smoother |
| $w_i$ | Weight $i$ in a regression model |
| $\mathbf{w}_k$ | Vector of weights at the time step $k$ in a regression model |
| $\mathbf{w}(t)$ | Gaussian white noise process |
| $W$ | Weight in the cubature or unscented approximation |

| | |
|---|---|
| $W_i$ | $i$th weight in sigma-point approximation or in Gauss–Hermite quadrature |
| $W_i^{(\mathrm{m})}$ | Mean weight of the unscented transform |
| $W_i^{(\mathrm{m})\prime}$ | Mean weight of the unscented transform |
| $W_i^{(\mathrm{c})}$ | Covariance weight of the unscented transform |
| $W_i^{(\mathrm{c})\prime}$ | Covariance weight of the unscented transform |
| $W_{i_1,\ldots,i_n}$ | Weight in multivariate Gauss–Hermite cubature |
| $x$ | Scalar random variable or state, sometimes regressor variable, or a generic scalar variable |
| $\mathbf{x}$ | Random variable or state |
| $\mathbf{x}^{(i)}$ | $i$th Monte Carlo draw from the distribution of $\mathbf{x}$ |
| $\mathbf{x}_k$ | State at the time step $k$ |
| $\tilde{\mathbf{x}}_k$ | Augmented state at the time step $k$ |
| $\mathbf{x}_{0:k}$ | Set containing the state vectors $\{\mathbf{x}_0, \ldots, \mathbf{x}_k\}$ |
| $\mathbf{x}_{0:k}^{(i)}$ | The history of the states in the particle $i$ |
| $\tilde{\mathbf{x}}_{0:T}^{(j)}$ | State trajectory simulated by a backward-simulation particle smoother |
| $\mathbf{X}$ | Matrix of regressors |
| $\mathbf{X}_k$ | Matrix of regressors up to the time step $k$ |
| $\mathcal{X}^{(\cdot)}$ | Sigma point of $\mathbf{x}$ |
| $\tilde{\mathcal{X}}^{(\cdot)}$ | Augmented sigma point of $\mathbf{x}$ |
| $\mathcal{X}_k^{(\cdot)}$ | Sigma point of the state $\mathbf{x}_k$ |
| $\tilde{\mathcal{X}}_k^{(\cdot)}$ | Augmented sigma point of the state $\mathbf{x}_k$ |
| $\hat{\mathcal{X}}_k^{(\cdot)}$ | Predicted sigma point of the state $\mathbf{x}_k$ |
| $\mathcal{X}_k^{-(\cdot)}$ | Sigma point of the predicted state $\mathbf{x}_k$ |
| $\tilde{\mathcal{X}}_k^{-(\cdot)}$ | Augmented sigma point of the predicted state $\mathbf{x}_k$ |
| $\mathbf{y}$ | Random variable or measurement |
| $\mathbf{y}_k$ | Measurement at the time step $k$ |
| $\mathbf{y}_{1:k}$ | Set containing the measurement vectors $\{\mathbf{y}_1, \ldots, \mathbf{y}_k\}$ |
| $\mathcal{Y}^{(\cdot)}$ | Sigma point of $\mathbf{y}$ |
| $\tilde{\mathcal{Y}}^{(\cdot)}$ | Augmented sigma point of $\mathbf{y}$ |
| $\hat{\mathcal{Y}}_k^{(\cdot)}$ | $i$th predicted sigma point of the measurement $\mathbf{y}_k$ at step $k$ |
| $Z$ | Normalization constant |
| $Z_k$ | Normalization constant at the time step $k$ |
| $\infty$ | Infinity |

**Abbreviations**

| | |
|---|---|
| ADF | Assumed density filter |
| AM | Adaptive Metropolis (algorithm) |
| AMCMC | Adaptive Markov chain Monte Carlo |
| AR | Autoregressive (model) |
| ARMA | Autoregressive moving average (model) |
| ASIR | Auxiliary sequential importance resampling |
| BS-PS | Backward-simulation particle smoother |
| CDKF | Central differences Kalman filter |
| CKF | Cubature Kalman filter |
| CLT | Central limit theorem |
| CPF | Cubature particle filter |
| CRLB | Cramér–Rao lower bound |
| DLM | Dynamic linear model |
| DOT | Diffuse optical tomography |
| DSP | Digital signal processing |
| EC | Expectation correction |
| EEG | Electroencephalography |
| EKF | Extended Kalman filter |
| EM | Expectation–maximization |
| EP | Expectation propagation |
| ERTSS | Extended Rauch–Tung–Striebel smoother |
| FHKF | Fourier–Hermite Kalman filter |
| FHRTSS | Fourier–Hermite Rauch–Tung–Striebel smoother |
| fMRI | Functional magnetic resonance imaging |
| GHKF | Gauss–Hermite Kalman filter |
| GHPF | Gauss–Hermite particle filter |
| GHRTSS | Gauss–Hermite Rauch–Tung–Striebel smoother |
| GPB | Generalized pseudo-Bayesian |
| GPS | Global positioning system |
| HMC | Hamiltonian (or hybrid) Monte Carlo |
| HMM | Hidden Markov model |
| IMM | Interacting multiple model (algorithm) |
| INS | Inertial navigation system |
| IS | Importance sampling |
| InI | Inverse imaging |
| KF | Kalman filter |
| LMS | Least mean squares |
| LQG | Linear quadratic Gaussian (regulator) |

| | |
|---|---|
| LS | Least squares |
| MA | Moving average (model) |
| MAP | Maximum a posteriori |
| MC | Monte Carlo |
| MCMC | Markov chain Monte Carlo |
| MEG | Magnetoencephalography |
| MH | Metropolis–Hastings |
| MKF | Mixture Kalman filter |
| ML | Maximum likelihood |
| MLP | Multi-layer perceptron |
| MMSE | Minimum mean squared error |
| MNE | Minimum norm estimate |
| MSE | Mean squared error |
| PF | Particle filter |
| PMCMC | Particle Markov chain Monte Carlo |
| PMMH | Particle marginal Metropolis–Hastings |
| PS | Particle smoother |
| QKF | Quadrature Kalman filter |
| RAM | Robust adaptive Metropolis (algorithm) |
| RBPF | Rao–Blackwellized particle filter |
| RBPS | Rao–Blackwellized particle smoother |
| RMSE | Root mean squared error |
| RTS | Rauch–Tung–Striebel |
| RTSS | Rauch–Tung–Striebel smoother |
| SDE | Stochastic differential equation |
| SIR | Sequential importance resampling |
| SIR-PS | Sequential importance resampling particle smoother |
| SIS | Sequential importance sampling |
| SLDS | Switching linear dynamic system |
| SLF | Statistically linearized filter |
| SLRTSS | Statistically linearized Rauch–Tung–Striebel smoother |
| SMC | Sequential Monte Carlo |
| TVAR | Time-varying autoregressive (model) |
| UKF | Unscented Kalman filter |
| UPF | Unscented particle filter |
| URTSS | Unscented Rauch–Tung–Striebel smoother |
| UT | Unscented transform |