Information Fusion 32 (2016) 1–11

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Applying Gaussian distributed constraints to Gaussian distributed variables

Andrew W. Palmer*, Andrew J. Hill, Steven J. Scheding

Australian Centre for Field Robotics, The University of Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:
Received 5 March 2015
Revised 4 January 2016
Accepted 18 February 2016
Available online 27 February 2016

Keywords:
Constrained Kalman filter
Uncertain constraints

ABSTRACT

This paper develops an analytical method of truncating inequality constrained Gaussian distributed variables where the constraints are themselves described by Gaussian distributions. Existing truncation methods either assume hard constraints, or use numerical methods to handle uncertain constraints. The proposed approach introduces moment-based Gaussian approximations of the truncated distribution. This method can be applied to numerous problems, with the motivating problem being Kalman filtering with uncertain constraints. In a simulation example, the developed method is shown to outperform unconstrained Kalman filtering by over 40% and hard-constrained Kalman filtering by over 17%.

1. Introduction

Gaussian distributioned variables are widely used to represent the state of a system in many problems ranging from state estimation [1] to scheduling [2,3]. In practice, the state vectors in many systems are known to satisfy inequality constraints. Examples of state-constrained systems include health monitoring [4], vision systems [5], robotics [6], binary sensor networks [7], and object tracking [8]. This paper deals specifically with systems that are subject to inequality constraints where the constraints themselves have uncertainty described by Gaussian distributions. Constraints described by Gaussian distributions can arise from many sources in state estimation problems including discrete sensors, such as position or level switches, that have uncertainty on their activation point, obstacles whose positions are uncertain, and other physical and model-derived bounds such as maximum fuel levels based on historical fuel burn rates. Constrained Gaussian distributed variables also appear in scheduling applications where the distribution describing the time at which an event is predicted to occur is constrained by the time distributions of other events.

Hard inequality constraints are well studied [1], where the main approaches are estimate projection [4], gain projection [9], and Probability Density Function (PDF) truncation [10]. Estimate and gain projection approaches incorporate the constraints into the derivation of the Kalman filter, resulting in a constrained optimisation problem that can be solved using quadratic programming, least squares approaches, amongst others [1,11]. Truncation methods, on the other hand, are applied directly to the PDF resulting from a Kalman filter, as outlined in Fig. 1. This approach truncates the PDF at the constraints and calculates the mean and covariance of the truncated PDF, which become the constrained state estimate and its covariance. The PDF truncation approach was shown in [10] to, in general, outperform the estimate projection method. The truncation approach has been applied to probabilistic collision checking for robots [12], and has been extended to non-linear systems [13,14].

Soft constraints correspond to uncertain or noisy constraints, and are less studied than hard constraints. Soft equality constraints are typically incorporated as noisy measurements [1,15]. However, soft inequality constraints are significantly more difficult to deal with, and numerical filters such as a Particle Filter (PF) are typically used for these problems [16]. Several numerical methods have been examined for incorporating soft constraints into the Kalman filter. A numerical PDF truncation method was used in [6] for robot localisation using Radio Frequency IDentification (RFID) tags, where the noise on the inequality constraints was highly non-Gaussian. Compared with a PF approach, the numerical PDF truncation method was 2 to 3 orders of magnitude faster while, in general, providing similar results. A similar RFID problem was examined in [7] where aspects of the Unscented Kalman Filter (UKF) and PF were combined—the prediction step used the standard UKF step, while the correction step was modified to weight the sigma-points of the UKF in a similar manner to the weighting process in a PF. It was shown to outperform a PF as well as the Quantised Extended Kalman Filter (QEKF) presented in [17].

* Corresponding author.
E-mail addresses: a.palmer@acfr.usyd.edu.au (A.W. Palmer), a.hill@acfr.usyd.edu.au (A.J. Hill), s.scheding@acfr.usyd.edu.au (S.J. Scheding).

http://dx.doi.org/10.1016/j.inffus.2016.02.008
1566-2535/

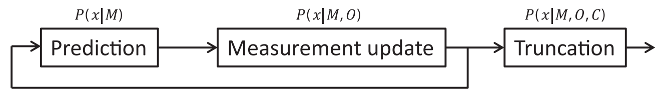

Fig. 1. The Kalman filter is run independent of the truncation method, with the truncation being applied to the state estimate that is the output of the Kalman filter. The prediction step of the Kalman filter results in a probability distribution describing the state, x, conditioned on the system model, M. The measurement update step further conditions the state estimate on the observations, O. Finally, the truncation step conditions the estimate on the constraints acting on the state, C.

The literature on soft inequality constraints has focused on constraints with non-Gaussian distributions, where the constrained state estimates are, by necessity, calculated using numerical methods. The main contribution of this paper is an analytical method for PDF truncation with soft constraints where the soft constraints are described by Gaussian distributions. This reduces the computational requirement compared to numerical methods, and it is shown to provide superior estimation performance compared to unconstrained and hard-constrained state estimation methods. The truncation approach presented in this paper is not limited to Kalman filters and can be applied to any constrained system using Gaussian distributions to represent the state and constraints.

The rest of this paper is structured as follows: Section 2 introduces the constrained Kalman filtering problem, Section 3 shows how the state and constraints can be transformed such that each state has only one constraint acting on it, Section 4 presents the truncation method for a one-sided constraint, and Section 5 extends this to an interval constraint. The performance of the methods are evaluated in Section 6, and the paper is concluded in Section 7. Appendix A and Appendix B provide in-depth derivations of the integrals used in this paper.

2. Problem definition

This paper adapts the notation used in [10]. A discrete linear time-invariant system is described by:

$$\boldsymbol{x}(k) = \boldsymbol{F}\boldsymbol{x}(k-1) + \boldsymbol{G}\boldsymbol{u}(k) + \boldsymbol{w}(k)$$
$$\boldsymbol{y}(k) = \boldsymbol{H}\boldsymbol{x}(k) + \boldsymbol{v}(k) \tag{1}$$

where k is the time index, $\boldsymbol{x}$ is the state vector with n states, $\boldsymbol{u}$ is the vector of known control inputs, and $\boldsymbol{y}$ is the vector of measurements. The vectors $\boldsymbol{w}$ and $\boldsymbol{v}$ contain the process and measurement noise respectively. The process noise, $\boldsymbol{w}$, is assumed to be zero mean Gaussian white noise with a covariance matrix of $\boldsymbol{Q}$. The measurement noise, $\boldsymbol{v}$, is similarly assumed to be zero mean Gaussian white noise with a covariance matrix of $\boldsymbol{R}$. The noises at each time-step are assumed to be independent.

For the given system, the Kalman filter prediction equations are [18]:

$$\hat{\boldsymbol{x}}(k|k-1) = \boldsymbol{F}\hat{\boldsymbol{x}}(k-1|k-1) + \boldsymbol{G}\boldsymbol{u}(k-1)$$
$$\boldsymbol{P}(k|k-1) = \boldsymbol{F}\boldsymbol{P}(k-1|k-1)\boldsymbol{F}^T + \boldsymbol{Q} \tag{2}$$

and the measurement update equations are:

$$\boldsymbol{K} = \boldsymbol{P}(k|k-1)\boldsymbol{H}^T\left(\boldsymbol{H}\boldsymbol{P}(k|k-1)\boldsymbol{H}^T + \boldsymbol{R}\right)^{-1}$$
$$\hat{\boldsymbol{x}}(k|k) = \hat{\boldsymbol{x}}(k|k-1) + \boldsymbol{K}\left(\boldsymbol{y}(k) - \boldsymbol{H}\hat{\boldsymbol{x}}(k|k-1)\right)$$
$$\boldsymbol{P}(k|k) = \boldsymbol{P}(k|k-1) - \boldsymbol{K}\boldsymbol{H}\boldsymbol{P}(k|k-1) \tag{3}$$

where $\hat{\boldsymbol{x}}(k|k)$ is the state estimate, and $\boldsymbol{P}(k|k)$ is the covariance of the state estimate. The state estimate is initialised with $\hat{\boldsymbol{x}}(0) = E[\boldsymbol{x}(0)]$, where $E[.]$ is the expectation operator. The covariance matrix is initialised with $\boldsymbol{P}(0) = E[(\boldsymbol{x}(0) - \hat{\boldsymbol{x}}(0))(\boldsymbol{x}(0) - \hat{\boldsymbol{x}}(0))^T]$.

Now consider the following s linearly independent constraints on the system:

$$A_m(k) \le \boldsymbol{\phi}_m^T(k)x(k) \le B_m(k) \qquad m = 1, \dots, s \tag{4}$$

where the constraints are uncertain and normally distributed:

$$A_m(k) \sim \mathcal{N}(\mu_{a,m}, \sigma_{a,m}^2) \qquad B_m(k) \sim \mathcal{N}(\mu_{b,m}, \sigma_{b,m}^2) \tag{5}$$

Eq. (4) describes a two-sided constraint on the linear function of the state described by $\boldsymbol{\phi}_m^T(k)\boldsymbol{x}(k)$. One sided constraints can be represented by setting $\mu_{a,m} = -\infty$, or $\mu_{b,m} = \infty$, and hard constraints can be implemented by setting $\sigma_{a,m} \approx 0$ or $\sigma_{b,m} \approx 0$ as required.

Given an estimate $\hat{\boldsymbol{x}}(k)$ with covariance $\boldsymbol{P}(k)$ at time k, the problem is to truncate the Gaussian PDF $\mathcal{N}(\hat{\boldsymbol{x}}(k), \boldsymbol{P}(k))$ using the s constraints described above, and then find the mean $\tilde{\boldsymbol{x}}(k)$ and covariance $\tilde{\boldsymbol{P}}(k)$ of the truncated PDF. The calculated mean and covariance represent the constrained estimate of the state.

3. Transforming the state vector and constraints

To apply the constraints via the truncation method, the state vector must be transformed so that the constraints are decoupled. This will result in s transformed constraints that each involve only one element of the transformed state, allowing the constraints to be enforced individually on each element of the transformed state. It should be noted that the order in which constraints are applied can change the final state estimate. However, if the initial constraints are decoupled, the order of constraint application does not change the result [10].

The transformation process is outlined in [1] and [10], and is summarised here in Eqs. (6)–(12) and (24)–(26). For ease of notation, the (k) after each variable will be dropped. Let the vector $\tilde{\boldsymbol{x}}_i$ be the truncated state estimate, and the matrix $\tilde{\boldsymbol{P}}_i$ be the covariance of $\tilde{\boldsymbol{x}}_i$, after the first $i-1$ constraints have been enforced. To initialise the process:

$$i = 1 \quad \tilde{\boldsymbol{x}}_i = \hat{\boldsymbol{x}} \quad \tilde{\boldsymbol{P}}_i = \boldsymbol{P} \tag{6}$$

The transformed state vector is given by:

$$\boldsymbol{z}_i = \boldsymbol{\rho}_i\boldsymbol{W}_i^{-1/2}\boldsymbol{T}_i^T(\boldsymbol{x} - \tilde{\boldsymbol{x}}_i) \tag{7}$$

where the matrices $\boldsymbol{T}_i$ and $\boldsymbol{W}_i$ are derived from the Jordan canonical decomposition of $\tilde{\boldsymbol{P}}_i$:

$$\boldsymbol{T}_i\boldsymbol{W}_i\boldsymbol{T}_i^T = \tilde{\boldsymbol{P}}_i \tag{8}$$

$\boldsymbol{T}_i$ is an orthogonal matrix, and $\boldsymbol{W}_i$ is a diagonal matrix. The matrix $\boldsymbol{\rho}_i$ is derived by the Gram–Schmidt orthogonalisation [19] which finds the orthogonal $\boldsymbol{\rho}_i$ that satisfies:

$$\boldsymbol{\rho}_i\boldsymbol{W}_i^{1/2}\boldsymbol{T}_i^T\boldsymbol{\phi}_i = \left[(\boldsymbol{\phi}_i^T\tilde{\boldsymbol{P}}_i\boldsymbol{\phi}_i)^{1/2} \quad 0 \quad \dots \quad 0\right]^T \tag{9}$$

Now only one element of $\boldsymbol{z}_i$ is constrained, and the states in the transformed state vector $\boldsymbol{z}_i$ are independent standard normal distributions. Let $\boldsymbol{e}_i$ be the ith column of an $n \times n$ identity matrix. Transforming the constraints results in:

$$C_i \le \boldsymbol{e}_i^T\boldsymbol{z}_i \le D_i \tag{10}$$

where

$$C_i \sim \mathcal{N}(\mu_{c,i}, \sigma_{c,i}^2)$$
$$\mu_{c,i} = \frac{\mu_{a,i} - \boldsymbol{\phi}_i^T\tilde{\boldsymbol{x}}_i}{\sqrt{\boldsymbol{\phi}_i^T\tilde{\boldsymbol{P}}_i\boldsymbol{\phi}_i}} \quad \sigma_{c,i} = \frac{\sigma_{a,i}}{\sqrt{\boldsymbol{\phi}_i^T\tilde{\boldsymbol{P}}_i\boldsymbol{\phi}_i}} \tag{11}$$

and

$$D_i \sim \mathcal{N}(\mu_{d,i}, \sigma_{d,i}^2)$$
$$\mu_{d,i} = \frac{\mu_{b,i} - \boldsymbol{\phi}_i^T\tilde{\boldsymbol{x}}_i}{\sqrt{\boldsymbol{\phi}_i^T\tilde{\boldsymbol{P}}_i\boldsymbol{\phi}_i}} \quad \sigma_{d,i} = \frac{\sigma_{b,i}}{\sqrt{\boldsymbol{\phi}_i^T\tilde{\boldsymbol{P}}_i\boldsymbol{\phi}_i}} \tag{12}$$

The equations for calculating the standard deviation of each constraint are not present in [1,10], but they are a trivial extension from the equations provided for calculating the mean.