Some notes on Extended Kalman Filters for Spacecraft Attitude Estimation

Rich Wareham

1 Introduction

These notes provide a heavily abbreviated summary of the use of the extended Kalman filter and modified Rodrigues rotation parametrisation in spacecraft attitude estimation. More detailed information can be found in [1, 2] but it is hoped that this note provides a concise reference for the concepts contained within those papers.

We begin with a review of two common parametrisations for rotations in $\mathbb{R}^3$, quaternions and the modified Rodrigues vector. We then define the state vector used to represent the current state of the spacecraft and which is updated by the Kalman filter. The various sensor measurement functions are then defined and finally we outline the complete Kalman filter algorithm.

2 Review of quaternions

The set of quaternions is conventionally denoted as $\mathbb{Q}$. This set can be associated with analogues of multiplication and addition to form a ring. A convenient parametrisation of a given quaternion is in terms of a *scalar* or "real" part and three independent *vector* or "imaginary" parts:

$$q \equiv q_1 + q_i \mathbf{i} + q_j \mathbf{j} + q_k \mathbf{k}$$

where $q_{\{q,i,j,k\}} \in \mathbb{R}$ and

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1.$$

A unit quaternion, q, is one where $q^2 = 1$. Quaternion conjugation is defined analogously to complex numbers:

$$(q_1 + q_i \mathbf{i} + q_j \mathbf{j} + q_k \mathbf{k})^* \triangleq q_1 - q_i \mathbf{i} - q_j \mathbf{j} - q_k \mathbf{k}$$

It should be clear from the above that both $\mathbb{R}$ and $\mathbb{C}$ are sub-groups of $\mathbb{Q}$.

We state without proof that the unit quaternion may be interpreted as representing a rotation in $\mathbb{R}^3$. Identifying $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ with some Cartesian basis of $\mathbb{R}^3$ we may bijectively map any vector $v \in \mathbb{R}^3$ to $\mathbb{Q}$. In so doing, the transformation

$$v' \leftarrow q \, v \, q^*$$

with $q^2 = -1$ represents some rotation of v. Letting $\bar{n} = n_i \mathbf{i} + n_j \mathbf{jk} + n_k \mathbf{k}$ be the unit vector pointing along the axis of rotation and θ be the rotation angle, we find that the quaternion representing the equivalent rotation is given by

$$q = \cos{(\theta/2)} + \sin{(\theta/2)} \, \bar{n}.$$

3   Review of the modified Rodrigues parameterisation

Below, we will be using the *Modified Rodrigues Parameters* representation. Letting $\langle q \rangle_s$ and $\langle q \rangle_v \equiv \langle q \rangle_i \, \mathbf{i} + \langle q \rangle_j \, \mathbf{j} + \langle q \rangle_k \, \mathbf{k}$ be the scalar (real) and vector (imaginary) components of the quaternion q, we have the MRP representation of q defined to be

$$\mathrm{mrp}(q) \triangleq \frac{1}{1 + \langle q \rangle_s} \, \langle q \rangle_v \equiv \bar{n} \tan\left(\frac{\theta}{4}\right)$$

where $\bar{n}$ is the unit vector pointing along the axis of rotation and θ is the angle of rotation. In particular, note that $\mathrm{mrp}(1) = \vec{0}$.

4   Representation of spacecraft state

In this section we outline the formation of the *state vector*, x_k, which we use to represent the state of the spacecraft at time instant k.

We firstly define two frames of reference. The *reference frame* is the absolute frame. Conventionally we may choose the z direction to be "up" and define x and y with respect to some horizontal reference. (We may choose x to be East and y to be North, for example.) The *body frame* is defined relative to the spacecraft. In this case z may be along the principal thrust axis and x and y may be defined with reference to some known point of reference on the spacecraft body.

Attitudes of spacecraft are usually represented as rotations mapping the reference frame to the spacecraft body frame. Define $q_k \in \mathbb{Q}$ to be the unit quaternion representing the rotation from the reference frame to the body frame at timestep k. Further let $\omega_k \in \mathbb{R}^3$ be the reference frame angular velocity vector of the spacecraft at timestep k.

We use the notation $\dot{x}$ to denote the differential of x with respect to time. Hence we will use $\ddot{p}_k$ to represent the linear acceleration of the spacecraft. For attitude estimation, it is assumed that this is simply the acceleration due to gravity.

There is an issue with using the attitude as represented by a quaternion directly in a Kalman filter. The interdependence of the four components of q_k is highly non-linear and, relatedly, the four components are a redundant parameterisation. We wish to use a quaternion representation, however, because it does not suffer from singularities of representation.

To work around this problem we shall instead keep track of the timestep-to-timestep *error* in attitude, $\Delta q_k \in \mathbb{Q}$. I.e., $q_k = \Delta q_k \, q_{k-1}$. Further we will represent Δq_k with the three-parameter MRP representation. As long as the error remains far from the representation's singular points there shouldn't be a problem. In the case of the MRP, this means that the spacecraft should not rotate more than $180°$ in one timestep.

We will be sensing orientation by integrating measurements from gyroscopes assumed to have some bias. Let b_k represent the gyro bias at timestep k such that the gyro measurement $\omega_k^{(g)}$ is related to the true body-frame angular velocity, $\omega_k^{(b)}$ by

$$\omega_k^{(g)} = \omega_k^{(b)} + b_k + \nu_k^{(g)}$$

where $\nu_k^{(g)}$ is the gyro noise at time k whose components are assumed to be i.i.d. samples from some zero-mean Gaussian process. The gyro bias, b_k, is assumed to be a Gaussian random walk. (I.e., $\dot{b}_k$ are drawn from some stationary zero-mean Gaussian process.)

In order to allow for gyro bias in our model, we shall estimate the instantaneous gyro bias b_k as part of the state vector.

The zero-noise kinematics of most of the state components may now be written directly:

$$\frac{\partial}{\partial t}\omega_k = \dot{\omega}_k, \quad \frac{\partial}{\partial t}\dot{\omega}_k = 0, \quad \frac{\partial}{\partial t}b_k = 0.$$

We shall split estimates relating to timestep k into *a priori* and *a posteriori* classes. Estimates which are *a priori* are formed from all measurements *prior* to timestep k whereas *a posteriori* estimates additionally take into account measurements from timestep k.

Our *a priori* state prediction step is therefore a simple Euler integration:

$$\hat{x}_{k|k-1} = f\left(\hat{x}_{k-1|k-1}; \Delta t\right) \triangleq \hat{x}_{k-1|k-1} + \Delta t \frac{\partial}{\partial t}\hat{x}_{k-1|k-1}$$

In addition we set $\mathrm{mrp}(\hat{\Delta}q_{k|k-1}) = \vec{0}$ to reflect our belief that the prediction error is zero and we also set $\hat{\dot{\omega}}_{k|k-1} = \vec{0}$ to reflect our belief that there is no imposed angular acceleration. In the case of active control, this last assumption should be replaced with the angular acceleration vector due to applied thrust.

The Jacobian of $f(\cdot)$ at $\hat{x}_{k-1|k-1}$ with respect to the state vector components, F_k, may be written down directly from the definition of the differential above.

For each time step, k, we wish to obtain the *a posteriori* attitude estimate $\hat{q}_{k|k}$. We can calculate it from the *a priori* and *a posteriori* state estimates $\hat{x}_{k|k-1}$ and $\hat{x}_{k|k}$. (Notice that in general we use $\hat{\ }$ to indicate the estimates of states and their components as opposed to their "true" values.) Define the *a priori* estimate of attitude to be

$$\hat{q}_{k|k-1} \triangleq q_{\hat{\omega}_{k-1|k-1}}^{\Delta t} \hat{q}_{k-1|k-1}$$

where $q_{\hat{\omega}_{k-1|k-1}}$ is a rotation around the axis aligned with $\hat{\omega}_{k-1|k-1}$. The angle of the rotation is the magnitude of $\hat{\omega}_{k-1|k-1}$.

The *a posteriori* estimate of attitude is obtained by applying the estimated error quaternion:

$$\hat{q}_{k|k} \triangleq \hat{\Delta}q_{k|k}\hat{q}_{k|k-1}.$$

5 Measurement functions

The accelerometer measurement function is

$$h_a(x_k; \hat{q}_{k|k-1}) = q'_k\, g\, q'^{*}_k$$

where g is the vector acceleration due to gravity in the reference frame and $q'_k \triangleq \hat{\Delta}q_{k|k-1}\,\hat{q}_{k|k-1}$.

The gyroscope measurement function is

$$h_g(x_k; \hat{q}_{k|k-1}) = q'_k\, \omega_k\, q'^{*}_k + b_k.$$

The magnetometer measurement function is

$$h_B(x_k; \hat{q}_{k|k-1}) = q'_k\, \frac{B}{|B|}\, q'^{*}_k$$

where B is the local reference-frame magnetic field vector.

Note that, in general, g and B are a function of p_k. For the sake of considering attitude-only motion, we assume that they are fixed and known.

For each measurement function we may also write down the Jacobian, H_k, centred on $\hat{x}_{k|k-1}$ with respect to the state vector components.

6   The extended Kalman filter

The extended Kalman filter (EKF) is a two stage estimation procedure. The *a priori* state estimate is formed from the *a posteriori* estimate from the previous time step. The *a posteriori* estimate for the current timestep is then computed by updating the *a priori* estimate.

In addition to the estimate of state, an estimate of the covariance of the state estimation error, P_k, is also maintained.

The Kalman filter assumes that the "true" state, x_k evolves according to:

$$x_k = f(x_{k-1}; q_{k-1}) + \nu_k^{(x)}$$

where $\nu_k^{(x)}$ are samples from some zero-mean Gaussian process with covariance Q_k. Similarly a measurement at timestep k, z_k is assumed to be given by

$$z_k = h(x_k) + \nu_k^{(z)}$$

where $h(\cdot)$ is the appropriate measurement function and $\nu_k^{(z)}$ are samples from a zero mean Gaussian process with covariance R_k. Each different sensor will have its own way of specifying R_k.

Letting F_k and H_k represent the Jacobians of $f(\cdot)$ and $h(\cdot)$ defined above, the EKF predict step is

$$\hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1}; \hat{q}_{k-1|k-1})$$
$$P_{k|k-1} = F_{k-1} P_{k-1|k-1} F_{k-1}^T + Q_k$$

where $P_{k|k-1}$ is the *a priori* prediction error covariance at timestep k, $P_{k-1|k-1}$ is the *a posteriori* estimation error covariance at timestep $k-1$ and Q_k is the process noise covariance for timestep $k-1$.

A sensor measurement, z_k, with associated measurement function, $h(\cdot)$, and sensor noise covariance R_k, is incorporated into the filter by way of the update step. First compute the innovation and innnovation covariance:

$$y_k = z_k - h(\hat{x}_{k|k-1}; \hat{q}_{k|k-1})$$
$$S_k = H_k P_{k|k-1} H_k^T + R_k.$$

Next we compute the Kalman gain:

$$K_k = P_{k|k-1} H_k^T S_k^{-1}.$$

This gain is used to compute the *a posteriori* state estimate mean and covariance:

$$P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1}$$
$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k y_k.$$

Multiple measurements at the same timestep may be handled by repeated updates after the prediction step. The relations for $\hat{q}_{k|k-1}$ and $\hat{q}_{k|k}$ were given above.

It is worth noting a curious property of $P_{k|k}$ in that it does not depend on the data. (Assuming that each sensor's measurement covariance is data-independent.) It follows that the state estimate covariance is entirely pre-determined.

References

[1] John L Crassidis and FL Markiey. Attitude estimation using modified rodrigues parameters, 1996.

[2] Christopher D Karlgaard and Hanspeter Schaub. Nonsingular attitude filtering using modified rodrigues parameters. *The Journal of the Astronautical Sciences*, 57(4):777–791, 2009.