Kalman Filter Tutorial for Balancing Robot

Rodrigo da Silva Guerra

April 3, 2008

1 General formulation

1.1 Process model

The equation 1 gives the internal dynamics of the process itself. The index k represent the discrete time steps. Usually the actual state variable x_k cannot be directly observed, but it can be estimated through some type of measurement device. The measurement device is modelled in equation 2. The measurement z_k can be directly observed. The column vectors w_k and v_k represent white noise added to the system. The Kalman filter presented in this tutorial will assume the system is modelled in accordance with these equations.

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \tag{1}$$

$$z_k = Hx_k + v_k \tag{2}$$

1.2 Noise covariances

When you have a single scalar stochastic variable you can estimate its variance. When you have a stochastic state vector then you don't have the variance anymore, but you have the covariance matrix instead. The covariance matrix is symmetric. The element at row i and column j of the covariance matrix tells the correlation between the elements at rows i and row j of the state vector. If the element at row i of x_k has no correlation at all with element at row j, then the element at row i and column j of the covariance matrix will be zero.

Below the equations 3 and 4 define the process and measurement covariance matrices, respectively.

$$Q = cov(w) \tag{3}$$

$$R = cov(v) \tag{4}$$

Usually the measurement noise covariance matrix R can be approximated or computed based on the characteristics of the sensor being used. Matrix R is more accessible than matrix Q and may even be estimated because we have direct access to the measurements. The matrix Q may be "hand tunned" by trial and error. Higher magnitudes in the elements of Q can be used to account for expected uncertainties (higher variance) between the corresponding elements of the state vector.

1.3 Kalman filter

The kalman filter is implemented in two steps:

1. Time update or prediction;
2. Measurement update or correction.

The time update (or prediction) is given by equations 5 and 6. The column vector $\hat{x}_k^-$ represents the estimation of the state variable x_k done *a priori*, before the actual measurement. The matrix P_k^- is an *a priori* estimation of the error covariance: $P_k^- = cov(x_k - \hat{x}_k^-)$.

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \tag{5}$$
$$P_k^- = AP_{k-1}A^T + Q \tag{6}$$

The measurement update (or correction) is given by equations 7, 8 and 9. The matrix K_k is the *Kalman gain* which is used in equation 8 to estimate the state variable by balancing between the predicted measurement and the actual measurement. The term $z_k - H\hat{x}_k^-$ is called *innovation*. The matrix P_k, computed in equation 9 is the error covariance at time k which will be used in order to predict the next error covariance in the next step, in equation 6.

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{7}$$
$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \tag{8}$$
$$P_k = (I - K_k H)P_k^- \tag{9}$$

The Kalman filtering algorithm consists basically of computing equations 5 and 6 (prediction equations) and then equations 7, 8 and 9 in a loop. For the initial iteration $\hat{P}_k^-$ can be initialized equals to Q or simply as I. In practice it should quickly converge so that its steady values can be sampled in a preliminary experiment and used for the next initialization.

2 Gyroscope + Accelerometer model

In this section specific implementations of the Kalman filter are introduced for the problem of estimating the tilt angle from gyroscope and accelerometer readings. Chip gyroscopes measure the angular rate (here noted as $\dot{\theta}_k^{gyro}$). This means the measurements have to be integrated over time in order to estimate the angle. Due to the gradual accumulation of small errors this causes a drift, introducing an incremental bias to the measurement. Chip 2-axis accelerometers can be used to measure the tilt angle directly: $\theta_k^{accel} = tan^{-1}(acc_x/axx_y)$ (in programming often there is a function called `atan2` which accounts for the sign of the arguments in order to return the angle in the correct quadrant.) Unfortunatelly accelerometers are very noisy and must be filtered, thus taking longer to update the correct angle.

In summary:

	Gyroscope	Accelerometer
Pros	Fast	No bias
Cons	Accumulate bias	Slow

2.1 Process model

The process is modeled following the same construction introduced in section 1.1. The variable θ_k represents the tilt angle (in rad) $\dot{\theta k}$ represents the angle increment (in rad/s) and $\dot{\delta}$ represents the increment bias error (also in rad/s). The constant dt represents the time elapsed between two measurements (in sec) which means $t_k = kdt$. The variable $\dot{\theta}_k^{gyro}$ represents the reading of the gyro (in rad/s) and θ_k^{accel} represents the reading of the accelerometer (in rad).

$$\begin{bmatrix} \theta_k \\ \dot{\theta}_k \\ \dot{\delta}_k \end{bmatrix} = \begin{bmatrix} 1 & 0 & -dt \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \theta_{k-1} \\ \dot{\theta}_{k-1} \\ \dot{\delta}_{k-1} \end{bmatrix} + \begin{bmatrix} dt & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\theta}_{k-1}^{gyro} \\ \theta_{k-1}^{accel} \end{bmatrix} + w_{k-1} \tag{10}$$

$$\begin{bmatrix} \dot{\theta}_k^{gyro} \\ \theta_k^{accel} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta_k \\ \dot{\theta}_k \\ \dot{\delta}_k \end{bmatrix} + v_k \tag{11}$$

Process noise covariance should be tunned experimentally. Here is an "initial guess":

$$Q = \begin{bmatrix} 0.2dt \\ 0.2 \\ 0.1 \end{bmatrix} \begin{bmatrix} 0.2dt \\ 0.2 \\ 0.1 \end{bmatrix}^T \tag{12}$$

The measurement noise covariance can be estimated by using some ground truth (e.g. attach the system to another device with an encoder). Another way to quickly estimate the measurement noise covariance is by consulting the variance of the measurement error in the datasheet or manual of the individual devices.

$$R = \begin{bmatrix} var(\dot{\theta}^{gyro}) & 0 \\ 0 & var(\theta^{accel}) \end{bmatrix} \tag{13}$$

Time update or prediction:

$$\begin{bmatrix} \hat{\theta}_k^- \\ \hat{\dot{\theta}}_k^- \\ \hat{\dot{\delta}}_k^- \end{bmatrix} = \begin{bmatrix} 1 & 0 & -dt \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\theta}_{k-1} \\ \hat{\dot{\theta}}_{k-1} \\ \hat{\dot{\delta}}_{k-1} \end{bmatrix} + \begin{bmatrix} dt & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\theta}_{k-1}^{gyro} \\ \theta_{k-1}^{accel} \end{bmatrix} \tag{14}$$

$$P_k^- = \begin{bmatrix} 1 & 0 & -dt \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix} P_{k-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ -dt & -1 & 1 \end{bmatrix} + Q \tag{15}$$

Measurement update or correction:

$$K_k = P_k^- \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} P_k^- \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} + R \right)^{-1} \tag{16}$$

$$\begin{bmatrix} \hat{\theta}_k \\ \hat{\dot{\theta}}_k \\ \hat{\dot{\delta}}_k \end{bmatrix} = \begin{bmatrix} \hat{\theta}_k^- \\ \hat{\dot{\theta}}_k^- \\ \hat{\dot{\delta}}_k^- \end{bmatrix} + K_k \left(\begin{bmatrix} \dot{\theta}_k^{gyro} \\ \theta_k^{accel} \end{bmatrix} - \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\theta}_k^- \\ \hat{\dot{\theta}}_k^- \\ \hat{\dot{\delta}}_k^- \end{bmatrix} \right) \tag{17}$$

$$P_k = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - K_k \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \right) P_k^- \tag{18}$$

Good luck!