Guide to Implementing Distributed Kalman Filter

11/2/2016

0 Preparation

For our filter, we require the following matrices and vectors:
$F \in R^{n\times n}$ - global model matrix (n is total number of state variables globally)
$G \in R^{n\times j}$ - state noise matrix (j is the number of noise sources)
$H \in R^{p\times n}$ - global observation matrix (p is the number of sensors observing the system)
$Q \in R^{j\times j}$ - process noise covariance
$R \in R^{p\times p}$ - measurement noise covariance
$S \in R^{n\times n}$ - state covariance
Refer to page 4921in the paper for more information.

We will split our global system into a number of subsystems which will independently run the distributed Kalman filtering algorithm. I will use the same example as provided in the paper on page 4923.
We are given the following state equations for the model as well as the observation vectors:

$$x_{k+1} = \begin{bmatrix} f_{11} & f_{12} & 0 & 0 & 0 \\ f_{21} & f_{22} & 0 & f_{24} & 0 \\ f_{31} & 0 & f_{33} & 0 & 0 \\ 0 & 0 & f_{43} & 0 & f_{45} \\ 0 & 0 & 0 & f_{54} & f_{55} \end{bmatrix} x_k + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & g_{32} \\ 0 & 0 \\ g_{51} & 0 \end{bmatrix} u_k = Fx_k + Gu_k \tag{1}$$

$$y_k = \begin{bmatrix} y_k^{(1)} \\ y_k^{(2)} \\ y_k^{(3)} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & 0 & 0 \\ 0 & h_{22} & h_{23} & h_{24} & 0 \\ 0 & 0 & 0 & h_{34} & h_{35} \end{bmatrix} x_k + \begin{bmatrix} w_k^{(1)} \\ w_k^{(2)} \\ w_k^{(3)} \end{bmatrix} = Hx_k + w_k \tag{2}$$

Note that we have 5 state variables total ($x_k \in R^5$), and 2 noise sources ($u_k \in R^2$). k represents each step in the Kalman filter procedure.

Observe matrix H. There are 3 sensors that make observations; the first sensor senses data for states 1,2, and 3, and operates on the data using h_{11}, h_{12}, h_{13}. The second sensor senses states 2,3,and 4, and uses h_{22}, h_{23}, h_{24}. Similarly the third sensor senses states 4 and 5. We will split our system according to the states that each sensor needs. In the example we will have 3 subsystems consisting of the sets of states (1,2,3), (2,3,4), and (4,5). See then that each row in H corresponds to a different subsystem, and the nonzero elements in each row correspond to the states that each subsystem will contain. Also note that subsystems can overlap and contain the same states.

With our system split up, we will need to identify matrices $F^{(l)}, G^{(l)}, H^{(l)}$, and $S^{(l)}$ for each subsystem l, as well as introduce a new matrix $D^{(l)}$ for each subsystem that will account for dependencies between subsystems.

For subsystem l, the new $F^{(l)}$ matrix will be a square matrix with elements extracted from the global F matrix corresponding to the states in the subsystem. For example, in subsystem 1 in the example, we have states 1,2, and 3. I take the f elements corresponding to these states, to get

$$F^{(1)} = \begin{bmatrix} f_{11} & f_{12} & 0 \\ f_{21} & f_{22} & 0 \\ f_{31} & 0 & f_{33} \end{bmatrix} \tag{3}$$

Similarly for subsystems 2 and 3, we have

$$F^{(2)} = \begin{bmatrix} f_{22} & 0 & f_{24} \\ 0 & f_{33} & 0 \\ 0 & f_{43} & 0 \end{bmatrix} \tag{4}$$

$$F^{(3)} = \begin{bmatrix} 0 & f_{45} \\ f_{54} & f_{55} \end{bmatrix} \tag{5}$$

To get the $G^{(l)}$ matrices, look at the global G matrix. For subsystem 1, no noise is added on to states 1 and 2 from either noise source, since the corresponding elements are 0. We see that noise from source 2 contributes to state 3, being multiplied by g_{32} first. Thus,

$$G^{(1)} = \begin{bmatrix} 0 \\ 0 \\ g_{32} \end{bmatrix} \tag{6}$$

and $u_k^{(1)} = u_{2,k}$ (the second noise source only). For subsystem 2, we once again only have the second noise source contributing to state 3, so

$$G^{(2)} = \begin{bmatrix} 0 \\ g_{32} \\ 0 \end{bmatrix} \tag{7}$$

$$u_k^{(2)} = u_{2,k} \tag{8}$$

Lastly for subsystem 3, noise source 1 contributes to state 5, so we have

$$G^{(3)} = \begin{bmatrix} 0 \\ g_{51} \end{bmatrix} \tag{9}$$

$$u_k^{(3)} = u_{1,k} \tag{10}$$

We need to introduce a new matrix/vector set D and d that will take care of relations between each subsystem. Refer back to the second row in global F: see that to calculate the second state, information from the fourth state in the previous timestep is needed (evidenced by the f_{24} term). The fourth state is not a part of subsystem 1 so this term has been neglected thus far. We include all such terms in matrix D, and vector d correspond to the states that correspond to the term. For example, in subsystem 1, we have

$$D^{(1)} = \begin{bmatrix} 0 \\ f_{24} \\ 0 \end{bmatrix} \tag{11}$$

$$d_k^{(1)} = x_{4,k} \tag{12}$$

The d vector is only a single term because only one outside state contributes to the calculation of states 1,2, and 3 in subsystem 1. For subsystem 2, we see that state 1 is needed in the calculation of states 2 and 3 in the next timestep, and state 5 is needed in calculating state 4 likewise. Thus we have

$$D^{(2)} = \begin{bmatrix} f_{21} & 0 \\ f_{31} & 0 \\ 0 & f_{45} \end{bmatrix} \tag{13}$$

$$d_k^{(2)} = \begin{bmatrix} x_{1,k} \\ x_{5,k} \end{bmatrix} \tag{14}$$

Lastly, for subsystem 3, we have

$$D^{(3)} = \begin{bmatrix} f_{43} \\ 0 \end{bmatrix} \tag{15}$$

$$d_k^{(3)} = x_{3,k} \tag{16}$$

The new state model equation for subsystem l in general is as follows:

$$x_{k+1}^{(l)} = F^{(l)} x_k^{(l)} + D^{(l)} d_k^{(l)} + G^{(l)} u_k^{(l)} \tag{17}$$

We do somewhat the same splitting procedure for the observation state equation. This will be somewhat simpler since we split our system based directly on the observation matrix H (one subsystem per row). Thus, the new state observation equation for subsystem l in general is

$$y_k^{(l)} = H^{(l)} x_k^{(l)} + w_k^{(l)} \tag{18}$$

Equation (2) shows $y_k^{(l)}$ and $w_k^{(l)}$, and $H^{(l)}$ is simply a row vector containing the nonzero elements in row l of matrix H (the elements corresponding to states in $x_k^{(l)}$. For example, for subsystem 1, $H^{(l)} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \end{bmatrix}$.

To obtain $S^{(l)}$ and $Q^{(l)}$ for subsystem l, simply take the elements in global matrices S and Q that correspond to the states in subsystem l (in the same fashion as how we extracted $F^{(l)}$ from F). For example, for S if we have

$$S = \begin{bmatrix} s_{11} & s_{12} & s_{13} & s_{14} & s_{15} \\ s_{21} & s_{22} & s_{23} & s_{24} & s_{25} \\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} \\ s_{41} & s_{42} & s_{43} & s_{44} & s_{45} \\ s_{51} & s_{52} & s_{53} & s_{54} & s_{55} \end{bmatrix} \tag{19}$$

then for each example subsystem, we have

$$S^{(1)} = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} \tag{20}$$

$$S^{(2)} = \begin{bmatrix} s_{22} & s_{23} & s_{24} \\ s_{32} & s_{33} & s_{34} \\ s_{42} & s_{43} & s_{44} \end{bmatrix} \tag{21}$$

$$S^{(3)} = \begin{bmatrix} s_{44} & s_{45} \\ s_{54} & s_{55} \end{bmatrix} \tag{22}$$

Finally, recall matrix R that represented the measurement noise covariance matrix for each sensor. We will assume that noise at each sensor is independent, so R is a diagonal matrix containing only variances. Each entry in the diagonal corresponds to a different sensor, and thus corresponds to a different subsystem. Later on, when I refer to $R^{(l)}$, I am referring to the lth element in the diagonal.

With this, we are finally able to begin the distributed Kalman filter process.

1 Initial conditions

Note that in the process of running the filter, we will be using information matrix Z which is the approximate inverse to S, and vector z which is an analogue to state vector x. There will be methods introduced to convert from Z to S and vice versa, as well as from z to x.

Start with the following (to be done for each subsystem):

$$z_{0|-1}^{(l)} = 0 \tag{23}$$
$$Z_{0|-1}^{(l)} = \text{lband\_inv}(S^{(l)}) \tag{24}$$

The lband_inv function is one that can be found in the github link. It takes 2 arguments; the matrix to invert as well as L, the size of the band of the approximate inverse output. For now, only use L = 1, that is, Z will be 1-banded (tridiagonal). Also, this is the only time that we will use the given $S^{(l)}$ matrices extracted from S; the Kalman filter will generate $Z^{(l)}$ matrices for later timesteps and thus we will calculate new $S^{(l)}$ matrices from that.

2 Observation fusion

Get sensor readings $y_k^{(l)}$ for subsystem l at timestep k and generate the following observation variables and matrix:

$$i_k^{(l)} = (H^{(l)})^T (R^{(l)})^{-1} y_k^{(l)} \tag{25}$$
$$\mathcal{I}^{(l)} = (H^{(l)})^T (R^{(l)})^{-1} H^{(l)} \tag{26}$$

Because the states in each subsystem overlap, we need to "fuse" our observation variables. To understand what this means, refer back to our example. Observe the global H matrix; we see that some states are read by multiple sensors (state 1 is only read by sensor 1 with the h_{11} term, state 2 is read by sensors 1 and 2 with h_{12} and h_{22} terms, and so on). Thus, for subsystem 1 that has states 1,2,and 3, some more information about states 2 and 3 are contained in the observation variables for subsystem 2, since states 2 and 3 also appear in subsystem 2. We will want to add this information to the observation variables for subsystem 1. Using our example, observe the following:

$$i^{(1)} = \begin{bmatrix} i_1^{(1)} \\ i_2^{(1)} \\ i_3^{(1)} \end{bmatrix}, i^{(2)} = \begin{bmatrix} i_2^{(2)} \\ i_3^{(2)} \\ i_4^{(2)} \end{bmatrix}, i^{(3)} = \begin{bmatrix} i_4^{(3)} \\ i_5^{(3)} \end{bmatrix} \tag{27}$$

where $i_m^{(l)}$ is the observation variable corresponding to state m in subsystem l. We fuse the information by simply adding information about the same states together, so our new observation variables for each subsystem are then

$$i_f^{(1)} = \begin{bmatrix} i_1^{(1)} \\ i_2^{(1)} + i_2^{(2)} \\ i_3^{(1)} + i_3^{(2)} \end{bmatrix}, i_f^{(2)} = \begin{bmatrix} i_2^{(2)} + i_2^{(1)} \\ i_3^{(2)} + i_3^{(1)} \\ i_4^{(2)} + i_4^{(3)} \end{bmatrix}, i_f^{(3)} = \begin{bmatrix} i_4^{(3)} + i_4^{(2)} \\ i_5^{(3)} \end{bmatrix} \tag{28}$$

similarly for the $\mathcal{I}^{(l)}$ matrices, if we have

$$\mathcal{I}^{(1)} = \begin{bmatrix} \mathcal{I}_{11}^{(1)} & \mathcal{I}_{12}^{(1)} & \mathcal{I}_{13}^{(1)} \\ \mathcal{I}_{21}^{(1)} & \mathcal{I}_{22}^{(1)} & \mathcal{I}_{23}^{(1)} \\ \mathcal{I}_{31}^{(1)} & \mathcal{I}_{32}^{(1)} & \mathcal{I}_{33}^{(1)} \end{bmatrix} \tag{29}$$
$$\mathcal{I}^{(2)} = \begin{bmatrix} \mathcal{I}_{22}^{(2)} & \mathcal{I}_{23}^{(2)} & \mathcal{I}_{24}^{(2)} \\ \mathcal{I}_{32}^{(2)} & \mathcal{I}_{33}^{(2)} & \mathcal{I}_{34}^{(2)} \\ \mathcal{I}_{42}^{(2)} & \mathcal{I}_{43}^{(2)} & \mathcal{I}_{44}^{(2)} \end{bmatrix} \tag{30}$$
$$\mathcal{I}^{(3)} = \begin{bmatrix} \mathcal{I}_{44}^{(3)} & \mathcal{I}_{45}^{(3)} \\ \mathcal{I}_{54}^{(3)} & \mathcal{I}_{55}^{(3)} \end{bmatrix} \tag{31}$$

then our fused matrices are as follows:

$$\mathcal{I}_f^{(1)} = \begin{bmatrix} \mathcal{I}_{11}^{(1)} & \mathcal{I}_{12}^{(1)} & \mathcal{I}_{13}^{(1)} \\ \mathcal{I}_{21}^{(1)} & \mathcal{I}_{22}^{(1)} + \mathcal{I}_{22}^{(2)} & \mathcal{I}_{23}^{(1)} + \mathcal{I}_{23}^{(2)} \\ \mathcal{I}_{31}^{(1)} & \mathcal{I}_{32}^{(1)} + \mathcal{I}_{32}^{(2)} & \mathcal{I}_{33}^{(1)} + \mathcal{I}_{33}^{(2)} \end{bmatrix} \tag{32}$$

$$\mathcal{I}_f^{(2)} = \begin{bmatrix} \mathcal{I}_{22}^{(2)} + \mathcal{I}_{22}^{(1)} & \mathcal{I}_{23}^{(2)} + \mathcal{I}_{23}^{(1)} & \mathcal{I}_{24}^{(2)} \\ \mathcal{I}_{32}^{(2)} + \mathcal{I}_{32}^{(1)} & \mathcal{I}_{33}^{(2)} + \mathcal{I}_{33}^{(1)} & \mathcal{I}_{34}^{(2)} \\ \mathcal{I}_{42}^{(2)} & \mathcal{I}_{43}^{(2)} & \mathcal{I}_{44}^{(2)} + \mathcal{I}_{44}^{(3)} \end{bmatrix} \tag{33}$$

$$\mathcal{I}_f^{(3)} = \begin{bmatrix} \mathcal{I}_{44}^{(3)} + \mathcal{I}_{44}^{(2)} & \mathcal{I}_{45}^{(3)} \\ \mathcal{I}_{54}^{(3)} & \mathcal{I}_{55}^{(3)} \end{bmatrix} \tag{34}$$

Find fused $i_k^{(l)}$ and $\mathcal{I}_k^{(l)}$ for every subsystem in this manner.

3 Local filter step

Run the filter step to update $Z_{k|k-1}^{(l)}$ to $Z_{k|k}^{(l)}$ and similarly for $z_{k|k-1}^{(l)}$:

$$Z_{k|k}^{(l)} = Z_{k|k-1}^{(l)} + \mathcal{I}_{f,k}^{(l)} \tag{35}$$

$$z_{k|k}^{(l)} = z_{k|k-1}^{(l)} + i_{f,k}^{(l)} \tag{36}$$

The notation $\mathcal{I}_{f,k}^{(l)}$ refers to the fused $\mathcal{I}$ matrices mentioned in the previous section, for timestep k and subsystem l. Likewise for $i_{f,k}^{(l)}$.

4 DICI-OR

Run the DICI-OR function as provided in the github code to convert $Z_{k|k}^{(l)}$ to $S_{k|k}^{(l)}$. The function takes 3 arguments: $Z_{k|k}^{(l)}$, $S_{k-1|k-1}^{(l)}$, and a parameter g that can be tweaked until we find an optimal value.

To convert $z_{k|k}^{(l)}$ to $x_{k|k}^{(l)}$, simply perform

$$x_{k|k}^{(l)} = S_{k|k}^{(l)} z_{k|k}^{(l)} \tag{37}$$

IMPORTANT: This $x_{k|k}^{(l)}$ is our state estimate that we will track for each timestep to see if it converges to the actual data.

5 Prepare for next iteration: obtaining $Z_{k|k-1}^{(l)}$ and $z_{k|k-1}^{(l)}$

The equation is

$$S_{k|k-1}^{(l)} = F^{(l)} S_{k-1|k-1}^{(l)} F^{(l)T} + F^{(l)} S_{k-1|k-1}^{x^{(l)}d^{(l)}} D^{(l)T} + (F^{(l)} S_{k-1|k-1}^{x^{(l)}d^{(l)}} D^{(l)T})^T \tag{38}$$

$$+ \; D^{(l)} S_{k-1|k-1}^{d^{(l)}d^{(l)}} D^{(l)T} + G^{(l)} Q^{(l)} G^{(l)T} \tag{39}$$

The details can be found on page 4930; we are still trying to figure out what the $S_{k-1|k-1}^{d^{(l)}d^{(l)}}$ and $S_{k-1|k-1}^{x^{(l)}d^{(l)}}$ matrices are.

Once we have $S_{k|k-1}^{(l)}$, run the lband_inv function once again to generate $Z_{k|k-1}^{(l)}$.

To get $z^{(l)}_{k|k-1}$, use the following:

$$z^{(l)}_{k|k-1} = Z^{(l)}_{k|k-1}(F^{(l)}x^{(l)}_{k-1|k-1} + D^{(l)}d^{(l)}_{k-1|k-1}) + f_1(Z^{(\mathcal{V})}_{k|k-1}, F^{(\mathcal{V})}, x^{(\mathcal{Q})}_{k-1|k-1}) \tag{40}$$

Details can be found on page 4931. We are still trying to figure out how to actually run this step (what $f_1(), \mathcal{V}, \mathcal{Q}$ are).