Two object tracking

SCJ Robertson

21st May, 2016

Introduction

This is an investigation of a two object tracking problem. The system is first represented as a Bayes Net, before a Junction Tree is constructed and the initial time steps of the system are investigated.

Writing anything I don't hate is nearly impossible, so this is mostly equations tied together with some string. The notation is still all over the place and there are probably a few copy and paste errors.

1. Representation

This section develops the model used in the tracking problem. Essentially the model is two Kalman Filters which have to choose which measurement they want and end up with a bit of both.

The proposed system is based on the following two rules:

1. There are always two targets, both of which are described by the same motion model.
2. At every time step, save for $t = 0$, each target must produce a measurement.

This may be obvious, but it should be concretely stated. It basically ensures we don't have to introduce any new targets or discard any existing targets, which is a huge simplification.

1.1 Bayes Net Representation

The Bayes Net for the system is shown in Figure 1.1, a_t is the association variable while the rest is pretty much the Kalman Filter representation we've seen before. The initial belief of each state is standard Kalman filter stuff:

$$p\left(\boldsymbol{x}_0^1\right) = \mathcal{N}\left(\boldsymbol{x}_0^1 | \boldsymbol{\mu}_0^1, \Sigma_0^1\right) \tag{1.1}$$

$$p\left(\boldsymbol{x}_0^2\right) = \mathcal{N}\left(\boldsymbol{x}_0^2 | \boldsymbol{\mu}_0^2, \Sigma_0^2\right) \tag{1.2}$$

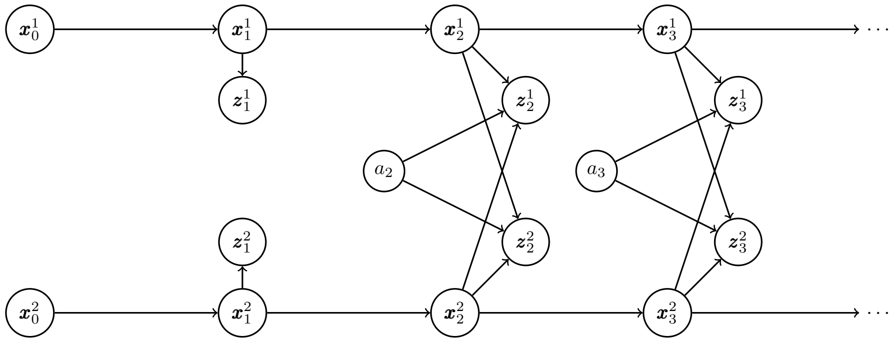


Figure 1.1: The initial frames for a Bayes Net modelling a two object tracking problem.

The prior's notation here is slightly misleading, we don't actually have specific knowledge about a target before the first measurements. All we know is that its launched on a golf course at some plausible velocity, that's hardly unique.

The belief of the current state of an object, $\boldsymbol{x}_t^i$, is given by the standard Kalman filter equations:

$$p\left(\boldsymbol{x}_t^1|\boldsymbol{x}_{t-1}^1\right) = \mathcal{N}\left(\boldsymbol{x}_t^1|A\boldsymbol{x}_{t-1}^1 + B\boldsymbol{u}, R\right) \tag{1.3}$$
$$p\left(\boldsymbol{x}_t^2|\boldsymbol{x}_{t-1}^2\right) = \mathcal{N}\left(\boldsymbol{x}_t^2|A\boldsymbol{x}_{t-1}^2 + B\boldsymbol{u}, R\right) \tag{1.4}$$

The initial measurements distributions, $\boldsymbol{z}_1^1$ and $\boldsymbol{z}_1^2$, are standard:

$$p\left(\boldsymbol{z}_1^1|\boldsymbol{x}_1^1\right) = \mathcal{N}\left(\boldsymbol{x}_1^1|C\boldsymbol{x}_1^1, Q\right) \tag{1.5}$$
$$p\left(\boldsymbol{z}_1^2|\boldsymbol{x}_1^2\right) = \mathcal{N}\left(\boldsymbol{x}_1^2|C\boldsymbol{x}_1^2, Q\right) \tag{1.6}$$

But measurement distributions for $t > 1$ are a bit different now:

$$p\left(\boldsymbol{z}_t^1|\boldsymbol{x}_t^1, \boldsymbol{x}_t^2, a_t\right) = \begin{cases} \mathcal{N}\left(\boldsymbol{z}_t^1|C\boldsymbol{x}_t^1, Q\right) \text{ if } a_t = 1 \\ \mathcal{N}\left(\boldsymbol{z}_t^1|C\boldsymbol{x}_t^2, Q\right) \text{ if } a_t = 2 \end{cases} \tag{1.7}$$

$$p\left(\boldsymbol{z}_t^2|\boldsymbol{x}_t^1, \boldsymbol{x}_t^2, a_t\right) = \begin{cases} \mathcal{N}\left(\boldsymbol{z}_t^2|C\boldsymbol{x}_t^2, Q\right) \text{ if } a_t = 1 \\ \mathcal{N}\left(\boldsymbol{z}_t^2|C\boldsymbol{x}_t^1, Q\right) \text{ if } a_t = 2 \end{cases} \tag{1.8}$$

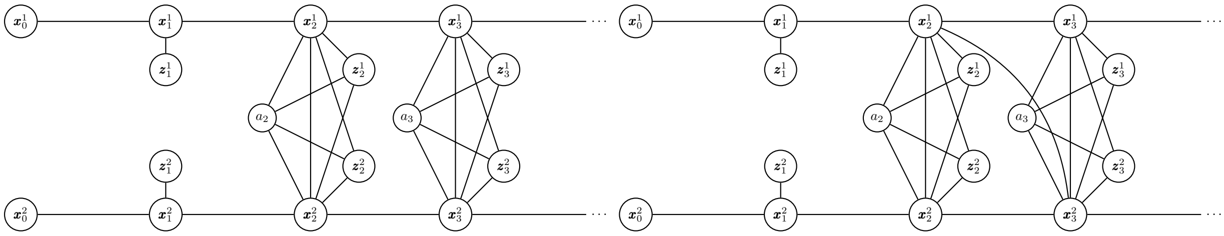
The association variable, a_t:

$$p\left(a_t\right) = \begin{cases} 0.5 \text{ if } a_t = 1 \\ 0.5 \text{ if } a_t = 2 \end{cases} \tag{1.9}$$

a_t is a binary random variable describing the probability of observation $\boldsymbol{z}_t^1$ being caused by $\boldsymbol{x}_t^1$ or $\boldsymbol{x}_t^2$. If we begin with a table enumerating all four possible causal combinations and assume both objects must create distinct measurements then $a_t = \{1, 2\}$ determines whether $\boldsymbol{z}_t^1$ was caused by $\boldsymbol{x}_t^1$ or $\boldsymbol{x}_t^2$. If $\boldsymbol{x}_t^1$ causes a certain measurement then the remaining measurement must be caused by $\boldsymbol{x}_t^2$.

1.2 Junction Tree Representation

An equivalent Junction tree will be created by applying the HUGIN algorithm to Figure 1.1. The undirected equivalent of the Bayes Net in Figure 1.1 is show in Figure 1.2a. After which the graph in Figure 1.2b is induced by variable elimination. The maximal cliques are then identified and directly used to construct Figure 1.3, this graph can then be rearranged into its final, more understandable representation in Figure 1.4.



(a) The moralised graph of Figure 1.1.

(b) The graph induced by variable elimination ordering $\boldsymbol{x}_0^1$, $\boldsymbol{x}_0^2$, $\boldsymbol{z}_1^1$, $\boldsymbol{z}_1^2$, $\boldsymbol{x}_1^2$, $\boldsymbol{x}_1^1$, $\boldsymbol{z}_2^1$, $\boldsymbol{z}_2^2$, a_2, $\boldsymbol{x}_2^2$, $\boldsymbol{x}_2^1$.

Figure 1.2: The graphs used in clique tree construction.

The cliques in Figure 1.4 are assigned the following distributions:

$$\boldsymbol{C}_{\boldsymbol{x}(0)}\left(\boldsymbol{x}_0^1, \boldsymbol{x}_0^2\right) = \mathcal{N}\left(\boldsymbol{x}_0^1 | \boldsymbol{\mu}_0^1, \Sigma_0^1\right) \mathcal{N}\left(\boldsymbol{x}_0^2 | \boldsymbol{\mu}_0^2, \Sigma_0^2\right) \tag{1.10}$$

$$\boldsymbol{C}_{\boldsymbol{z}(1)}\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2, \boldsymbol{z}_1^1, \boldsymbol{z}_1^2\right) = \mathcal{N}\left(\boldsymbol{z}_1^1 | C\boldsymbol{x}_1^1, Q\right) \mathcal{N}\left(\boldsymbol{z}_2^1 | C\boldsymbol{x}_2^1, Q\right) \tag{1.11}$$

$$\boldsymbol{C}_{\boldsymbol{x}(t)}\left(\boldsymbol{x}_{t-1}^1, \boldsymbol{x}_{t-1}^2, \boldsymbol{x}_t^1, \boldsymbol{x}_t^2\right) = \mathcal{N}\left(\boldsymbol{x}_t^1 | A\boldsymbol{x}_{t-1}^1 + B\boldsymbol{u}, R\right) \mathcal{N}\left(\boldsymbol{x}_t^2 | A\boldsymbol{x}_{t-1}^2 + B\boldsymbol{u}, R\right) \tag{1.12}$$

$$\begin{aligned} \boldsymbol{C}_{\boldsymbol{z}(t)}\left(\boldsymbol{x}_t^1, \boldsymbol{x}_t^2, \boldsymbol{z}_t^1, \boldsymbol{z}_t^2, a_t\right) &= p\left(\boldsymbol{z}_t^1 | \boldsymbol{x}_t^1, \boldsymbol{x}_t^2, a_t\right) p\left(\boldsymbol{z}_t^2 | \boldsymbol{x}_t^1, \boldsymbol{x}_t^2, a_t\right) p\left(a_t\right) \\ &= \begin{cases} \mathcal{N}\left(\boldsymbol{z}_t^1 | C\boldsymbol{x}_t^1, Q\right) \mathcal{N}\left(\boldsymbol{z}_t^2 | C\boldsymbol{x}_t^2, Q\right) \text{ if } a_t = 1 \\ \mathcal{N}\left(\boldsymbol{z}_t^1 | C\boldsymbol{x}_t^2, Q\right) \mathcal{N}\left(\boldsymbol{z}_t^2 | C\boldsymbol{x}_t^1, Q\right) \text{ if } a_t = 2 \end{cases} \end{aligned} \tag{1.13}$$

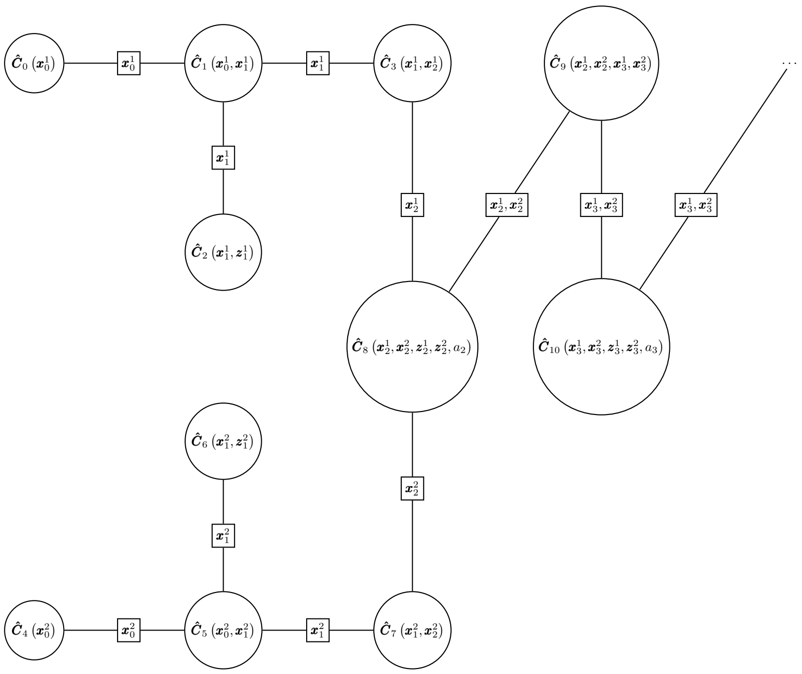


Figure 1.3: A Junction Tree constructed directly from the maximal cliques identified in Figure 1.2b.

2. Analysis

Now we investigate the proposed system, which is in general intractable, but will provide motivation for future developments.

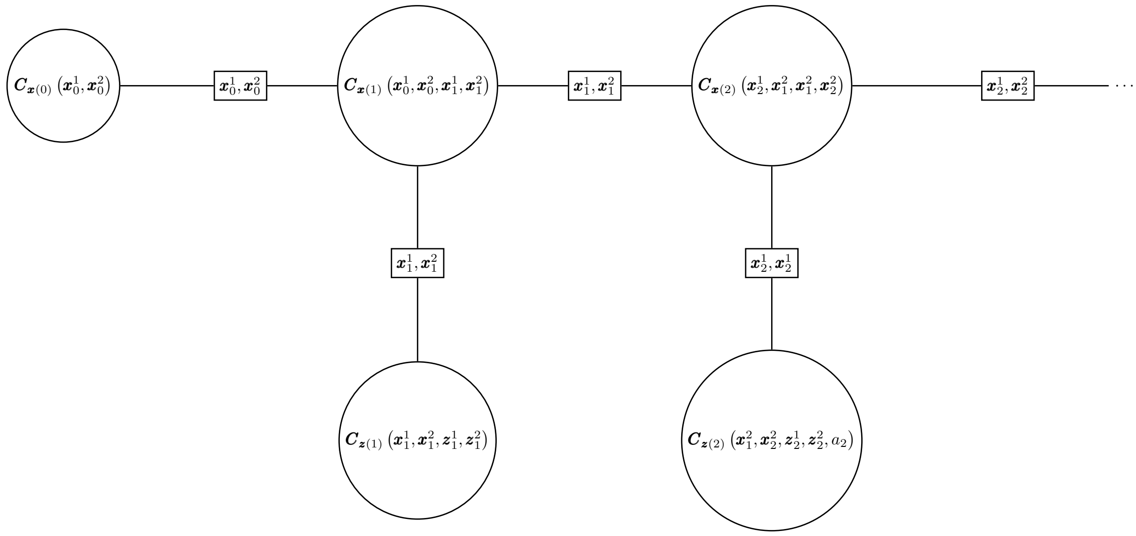


Figure 1.4: A better Junction Tree. This tree is equivalent to the Junction Tree in Figure 1.3, but easier to analyse.

2.1 Mathematical Investigation

This section is an investigation into the initial three frames[1] of the system. Figure 2.1 provides a useful analogue of what is happening in the initial frames, but it is the same for far too many reasons (mainly because drawing joint distributions in TIKZ is difficult).

2.1.1 The Measurement Message

All measurement distributions have the same form:

$$\boldsymbol{C}_{\boldsymbol{z}(t)}\left(\boldsymbol{x}_t^1, \boldsymbol{x}_t^2, \boldsymbol{z}_t^1, \boldsymbol{z}_t^2, a_t\right) = \begin{cases} \mathcal{N}\left(\boldsymbol{z}_t^1 | C\boldsymbol{x}_t^1, Q\right) \mathcal{N}\left(\boldsymbol{z}_t^2 | C\boldsymbol{x}_t^2, Q\right) \text{ if } a_t = 1 \\ \mathcal{N}\left(\boldsymbol{z}_t^1 | C\boldsymbol{x}_t^2, Q\right) \mathcal{N}\left(\boldsymbol{z}_t^2 | C\boldsymbol{x}_t^1, Q\right) \text{ if } a_t = 2 \end{cases} \tag{2.1}$$

$$= \begin{cases} \mathcal{C}\left(\boldsymbol{x}_t^1, \boldsymbol{z}_t^1; \hat{P}, \mathbf{0}, \hat{g}\right) \mathcal{C}\left(\boldsymbol{x}_t^2, \boldsymbol{z}_t^2; \hat{P}, \mathbf{0}, \hat{g}\right) \text{ if } a_t = 1 \\ \mathcal{C}\left(\boldsymbol{x}_t^1, \boldsymbol{z}_t^2; \hat{P}, \mathbf{0}, \hat{g}\right) \mathcal{C}\left(\boldsymbol{x}_t^2, \boldsymbol{z}_t^1; \hat{P}, \mathbf{0}, \hat{g}\right) \text{ if } a_t = 2 \end{cases} \tag{2.2}$$

Where,

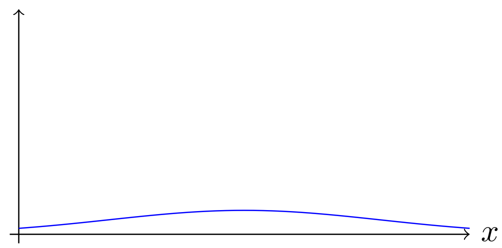
$$\hat{P} = \begin{bmatrix} C^T Q^{-1} C & -C^T Q^{-1} \\ -Q^{-1} C & Q^{-1} \end{bmatrix} \tag{2.3}$$

$$\hat{g} = -\ln\left((2\pi)^{k/2} |Q|^{1/2}\right) \tag{2.4}$$

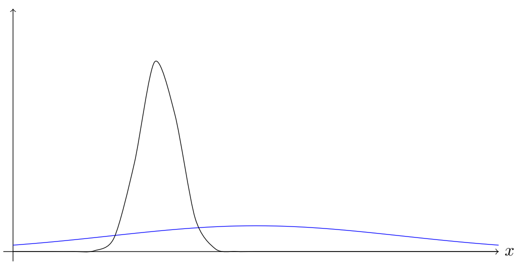
The measurement update message, after observing $\boldsymbol{z}_t^1$ and $\boldsymbol{z}_t^2$:

$$\begin{aligned} \delta_{\boldsymbol{z}(t)\to\boldsymbol{x}(t)}\left(\boldsymbol{x}_t^1, \boldsymbol{x}_t^2\right) &= \int \boldsymbol{C}_{\boldsymbol{z}(t)}\left(\boldsymbol{x}_t^1, \boldsymbol{x}_t^2, \boldsymbol{z}_t^1 = \boldsymbol{z}_t^2, \boldsymbol{z}_t^2 = \boldsymbol{z}_t^2, a_t\right) d\{a_t\} \\ &= \frac{1}{2}\mathcal{C}\left(\boldsymbol{x}_t^1, \boldsymbol{x}_t^2; P_{\boldsymbol{z}}, \boldsymbol{h}_{\boldsymbol{z}(1)}, g_{\boldsymbol{z}}\right) + \frac{1}{2}\mathcal{C}\left(\boldsymbol{x}_t^1, \boldsymbol{x}_t^2; P_{\boldsymbol{z}}, \boldsymbol{h}_{\boldsymbol{z}(2)}, g_{\boldsymbol{z}}\right) \end{aligned} \tag{2.5}$$

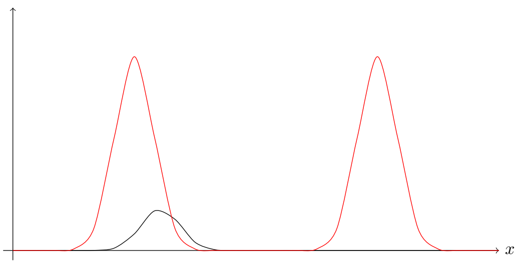
[1] Technically, two frames. Frame 1, $\boldsymbol{x}_0$, can be absorbed and then ignored.



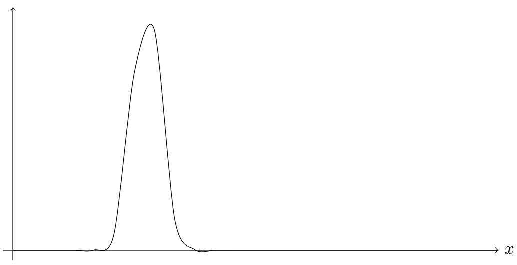
(a) As always, we initially know nothing about state x_1^1.



(b) The belief is updated with a specific measurement, the object now has a stronger association with a specific trajectory.



(c) Before the measurement update of x_2^1. The measurement update will always think both measurements were likely to been caused by x_1.



(d) After the measurement update, the trajectory associated with x_1 dominates, though this is technically still a mixture of Gaussians.

Figure 2.1: A system analogous to that developed in the Chapter 1. The distribution of x_t^1 is a Gaussian mixture with an exponential amount of components, a component for every possible trajectory introduced by the measurement update. This approach primes the system so that a target's state is strongly associated with a specific trajectory.

This a two component Gaussian mixture, where:

$$P_{\boldsymbol{z}} = \begin{bmatrix} C^T Q^{-1} C & 0 \\ 0 & C^T Q^{-1} C \end{bmatrix} \tag{2.6}$$

$$\boldsymbol{h}_{\boldsymbol{z}(1)} = \begin{bmatrix} C^T Q^{-1} \boldsymbol{z}_t^1 \\ C^T Q^{-1} \boldsymbol{z}_t^2 \end{bmatrix} \tag{2.7}$$

$$\boldsymbol{h}_{\boldsymbol{z}(1)} = \begin{bmatrix} C^T Q^{-1} \boldsymbol{z}_t^2 \\ C^T Q^{-1} \boldsymbol{z}_t^1 \end{bmatrix} \tag{2.8}$$

$$g_{\boldsymbol{z}} = -\ln\left((2\pi)^k |Q|\right) - \frac{1}{2}(\boldsymbol{z}_t^1)^T Q^{-1}(\boldsymbol{z}_t^1) - \frac{1}{2}(\boldsymbol{z}_t^2)^T Q^{-1}(\boldsymbol{z}_t^2) \tag{2.9}$$

2.1.2 Recursive Belief Message

At time step $t = 1$ the system is simply the product of two independent Kalman Filters. Rather than formal derivation, its easier just to propagate a Gaussian forward using the standard Kalman Filter equations:

$$\begin{aligned} \therefore \delta_{\boldsymbol{x}(1)\to\boldsymbol{x}(2)}\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2\right) &= \int \boldsymbol{C}_{\boldsymbol{x}(1)}\left(\boldsymbol{x}_0^1, \boldsymbol{x}_0^2, \boldsymbol{x}_1^1, \boldsymbol{x}_1^2\right) d\{\boldsymbol{x}_0^1, \boldsymbol{x}_0^2\} \\ &= \mathcal{C}\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2; P_{\boldsymbol{x}(1)}, \boldsymbol{h}_{\boldsymbol{x}(1)}, g_{\boldsymbol{x}(1)}\right) \end{aligned} \tag{2.10}$$

Where,

$$P_{\boldsymbol{x}(1)} = \begin{bmatrix} \left(\Sigma_1^1\right)^{-1} & 0 \\ 0 & \left(\Sigma_1^2\right)^{-1} \end{bmatrix} \tag{2.11}$$

$$\boldsymbol{h}_{\boldsymbol{x}(1)} = \begin{bmatrix} \left(\Sigma_1^1\right)^{-1} \boldsymbol{\mu}_1^1 \\ \left(\Sigma_1^2\right)^{-1} \boldsymbol{\mu}_1^2 \end{bmatrix} \tag{2.12}$$

$$\begin{aligned} g_{\boldsymbol{x}(1)} = &- \ln\left((2\pi)^{n/2} |\Sigma_1^1|^{1/2}\right) - \ln\left((2\pi)^{n/2} |\Sigma_1^2|^{1/2}\right) \\ &- \frac{1}{2}(\boldsymbol{\mu}_1^1)^T \left(\Sigma_1^1\right)^{-1} (\boldsymbol{\mu}_1^1) - \frac{1}{2}(\boldsymbol{\mu}_1^2)^T \left(\Sigma_1^2\right)^{-1} (\boldsymbol{\mu}_1^2) \end{aligned} \tag{2.13}$$

I never derived the scalar $g_{\boldsymbol{x}(t)}$ last time and most people seem to ignore it, but the outgoing message is Gaussian so $g_{\boldsymbol{x}(t)}$ should be valid.

2.1.3 The Second Frame

Before any updates occur, the following uninformed potential inhabits $\boldsymbol{C}_{\boldsymbol{x}(2)}$:

$$\boldsymbol{C}_{\boldsymbol{x}(2)}\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2, \boldsymbol{x}_2^1, \boldsymbol{x}_2^2\right) = \mathcal{C}_{\boldsymbol{x}(2)}\left(\boldsymbol{x}_1^1, \boldsymbol{x}_2^1, \boldsymbol{x}_1^2, \boldsymbol{x}_2^2; P_{\boldsymbol{x}(2)}, \boldsymbol{h}_{\boldsymbol{x}(2)}, g_{\boldsymbol{x}(2)}\right)$$

Where,

$$P_{\boldsymbol{x}(2)} = \begin{bmatrix} \begin{matrix} R^{-1} & -R^{-1}A \\ -A^T R^{-1} & A^T R^{-1} A \end{matrix} & 0 \\ 0 & \begin{matrix} R^{-1} & -R^{-1}A \\ -A^T R^{-1} & A^T R^{-1} A \end{matrix} \end{bmatrix} \tag{2.14}$$

$$\boldsymbol{h}_{\boldsymbol{x}(2)} = \begin{bmatrix} R^{-1}B\boldsymbol{u} \\ -AR^{-1}B\boldsymbol{u} \\ R^{-1}B\boldsymbol{u} \\ -AR^{-1}B\boldsymbol{u} \end{bmatrix} \tag{2.15}$$

$$g_{\boldsymbol{x}(2)} = -\ln((2\pi)^n |R|) - \frac{1}{2}\boldsymbol{h}_{\boldsymbol{x}(2)}^T P_{\boldsymbol{x}(2)}^{-1} \boldsymbol{h}_{\boldsymbol{x}(2)} \tag{2.16}$$

Belief Update

The outgoing message from $\boldsymbol{C}_{\boldsymbol{x}(2)}$:

$$\delta_{\boldsymbol{x}(2)\to\boldsymbol{x}(3)}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2\right) = \delta_{\boldsymbol{z}(2)\to\boldsymbol{x}(3)}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2\right) \underbrace{\int \boldsymbol{C}_{\boldsymbol{x}(2)}\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2, \boldsymbol{x}_2^1, \boldsymbol{x}_2^2\right) \delta_{\boldsymbol{x}(1)\to\boldsymbol{x}(2)}\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2\right) d\{\boldsymbol{x}_1^1, \boldsymbol{x}_1^2\}}_{\Psi\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2\right)}$$

$\Psi\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2\right)$ is nothing new, the established equations just have to be applied independently. Giving:

$$\Psi\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2\right) = \mathcal{C}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2; P_\Psi, \boldsymbol{h}_\Psi, g_\Psi\right) \tag{2.17}$$

Where,

$$
\begin{aligned}
P_\Psi &= \begin{bmatrix} \left(R + A\Sigma_1^1 A^T\right)^{-1} & 0 \\ 0 & \left(R + A\Sigma_1^2 A^T\right)^{-1} \end{bmatrix} \\
&= \begin{bmatrix} \left(\overline{\Sigma}_2^1\right)^{-1} & 0 \\ 0 & \left(\overline{\Sigma}_2^2\right)^{-1} \end{bmatrix} && (2.18) \\
\boldsymbol{h}_\Psi &= \begin{bmatrix} \left(R + A\Sigma_1^1 A^T\right)^{-1} \left(A\boldsymbol{\mu}_1^1 + B\boldsymbol{u}\right) \\ \left(R + A\Sigma_1^2 A^T\right)^{-1} \left(A\boldsymbol{\mu}_1^2 + B\boldsymbol{u}\right) \end{bmatrix} \\
&= \begin{bmatrix} \left(\overline{\Sigma}_2^1\right)^{-1} \overline{\boldsymbol{\mu}}_2^1 \\ \left(\overline{\Sigma}_2^1\right)^{-1} \overline{\boldsymbol{\mu}}_2^2 \end{bmatrix} && (2.19) \\
g_\Psi &= -\ln((2\pi)^{n/2}|R + A\Sigma_1^1 A^T|^{1/2}) - \ln((2\pi)^{n/2}|R + A\Sigma_1^1 A^T|^{1/2}) \\
&- \frac{1}{2}\boldsymbol{h}_\Psi^T P_\Psi^{-1} \boldsymbol{h}_\Psi && (2.20)
\end{aligned}
$$

g_Ψ is not necessarily correct, but I really hate dealing with it, so I'm not going to.

Measurement Update

Eventually,

$$
\begin{aligned}
\delta_{\boldsymbol{x}(2)\to\boldsymbol{x}(3)}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2\right) &= \delta_{\boldsymbol{z}(2)\to\boldsymbol{x}(2)}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2\right) \Psi\left(\boldsymbol{x}_1^1, \boldsymbol{x}_1^2\right) \\
&= \mathcal{C}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2; P_{\boldsymbol{z}(2)}, h_{\boldsymbol{z}(2)}, g_{\boldsymbol{z}(2)}\right) \mathcal{C}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2; P_\Psi, \boldsymbol{h}_\Psi, g_\Psi\right) \\
&= \frac{1}{2}\mathcal{C}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2; P_{\boldsymbol{z}}, \boldsymbol{h}_{\boldsymbol{z}(1)}, g_{\boldsymbol{z}}\right) \mathcal{C}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2; P_\Psi, \boldsymbol{h}_\Psi, g_\Psi\right) \\
&+ \frac{1}{2}\mathcal{C}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2; P_{\boldsymbol{z}}, \boldsymbol{h}_{\boldsymbol{z}(2)}, g_{\boldsymbol{z}}\right) \mathcal{C}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2; P_\Psi, \boldsymbol{h}_\Psi, g_\Psi\right) \\
&= \frac{1}{2}\mathcal{C}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2; P_{\boldsymbol{x}(2)}, \boldsymbol{h}_{\boldsymbol{x}(2)}^1, g_{\boldsymbol{x}(2)}\right) + \frac{1}{2}\mathcal{C}\left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2; P_{\boldsymbol{x}(2)}, \boldsymbol{h}_{\boldsymbol{z}(2)}^2, g_{\boldsymbol{x}(2)}\right)
\end{aligned}
$$

Where,

$$
\begin{aligned}
P_{\boldsymbol{x}(2)} &= \begin{bmatrix} \left(\overline{\Sigma}_2^1\right)^{-1} + C^T Q^{-1} C & 0 \\ 0 & \left(\overline{\Sigma}_2^2\right)^{-1} + C^T Q^{-1} C \end{bmatrix} && (2.21) \\
\boldsymbol{h}_{\boldsymbol{x}(2)}^1 &= \begin{bmatrix} C^T Q^{-1} \boldsymbol{z}_2^1 + \left(\overline{\Sigma}_2^1\right)^{-1} \overline{\boldsymbol{\mu}}_2^1 \\ C^T Q^{-1} \boldsymbol{z}_2^2 + \left(\overline{\Sigma}_2^2\right)^{-1} \overline{\boldsymbol{\mu}}_2^2 \end{bmatrix} && (2.22) \\
\boldsymbol{h}_{\boldsymbol{x}(2)}^2 &= \begin{bmatrix} C^T Q^{-1} \boldsymbol{z}_2^2 + \left(\overline{\Sigma}_2^1\right)^{-1} \overline{\boldsymbol{\mu}}_2^1 \\ C^T Q^{-1} \boldsymbol{z}_2^1 + \left(\overline{\Sigma}_2^2\right)^{-1} \overline{\boldsymbol{\mu}}_2^2 \end{bmatrix} && (2.23)
\end{aligned}
$$

From this,

$$\Sigma_{\boldsymbol{x}(2)} = \begin{bmatrix} \left(I - K_2^1 C\right) \overline{\Sigma}_2^1 & 0 \\ 0 & \left(I - K_2^2 C\right) \overline{\Sigma}_2^2 \end{bmatrix} \tag{2.24}$$

$$\begin{aligned}
\boldsymbol{\mu}_{\boldsymbol{x}(2)}^1 &= P_{\boldsymbol{x}(2)} \boldsymbol{h}_{\boldsymbol{x}(2)}^1 \\
&= \begin{bmatrix} \overline{\boldsymbol{\mu}}_2^1 + K_1^1 \left(\overline{\boldsymbol{\mu}}_2^1 - C\boldsymbol{z}_2^1\right) \\ \overline{\boldsymbol{\mu}}_2^2 + K_1^2 \left(\overline{\boldsymbol{\mu}}_2^2 - C\boldsymbol{z}_2^2\right) \end{bmatrix} \\
\boldsymbol{\mu}_{\boldsymbol{x}(2)}^2 &= P_{\boldsymbol{x}(2)} \boldsymbol{h}_{\boldsymbol{x}(2)}^2 \\
&= \begin{bmatrix} \overline{\boldsymbol{\mu}}_2^1 + K_1^1 \left(\overline{\boldsymbol{\mu}}_2^1 - C\boldsymbol{z}_2^2\right) \\ \overline{\boldsymbol{\mu}}_2^2 + K_1^2 \left(\overline{\boldsymbol{\mu}}_2^2 - C\boldsymbol{z}_2^1\right) \end{bmatrix}
\end{aligned}$$

This is all good and well, but its difficult to see the state of a single target from the joint representation. Looking at just $\boldsymbol{x}_2^1$:

$$\begin{aligned}
\Phi\left(\boldsymbol{x}_2^1\right) &= \int \delta_{\boldsymbol{x}(2)\to\boldsymbol{x}(3)} \left(\boldsymbol{x}_2^1, \boldsymbol{x}_2^2\right) d\boldsymbol{x}_2^2 \\
&= \frac{1}{2}\mathcal{C}\left(\boldsymbol{x}_2^1; P_{\boldsymbol{x}}, \boldsymbol{h}_{\boldsymbol{x}}^1, g_{\boldsymbol{x}}\right) + \frac{1}{2}\mathcal{C}\left(\boldsymbol{x}_2^1; P_{\boldsymbol{x}}, \boldsymbol{h}_{\boldsymbol{x}}^2, g_{\boldsymbol{x}}\right)
\end{aligned} \tag{2.25}$$

Where,

$$P_{\boldsymbol{x}} = C^T Q^{-1} \boldsymbol{z}_2^1 + \left(\overline{\Sigma}_2^1\right)^{-1} \overline{\boldsymbol{\mu}}_2^1 \tag{2.26}$$

$$\boldsymbol{h}_{\boldsymbol{x}}^1 = C^T Q^{-1} \boldsymbol{z}_2^1 + \left(\overline{\Sigma}_2^1\right)^{-1} \overline{\boldsymbol{\mu}}_2^1 \tag{2.27}$$

$$\boldsymbol{h}_{\boldsymbol{x}}^2 = C^T Q^{-1} \boldsymbol{z}_2^2 + \left(\overline{\Sigma}_2^2\right)^{-1} \overline{\boldsymbol{\mu}}_2^2 \tag{2.28}$$

Now,

$$\Sigma_{\boldsymbol{x}(2)} = \left(I - K_1^1 C\right) \overline{\Sigma}_1^1 \tag{2.29}$$

$$\boldsymbol{\mu}_{\boldsymbol{x}(2)}^1 = \overline{\boldsymbol{\mu}}_2^1 + K_1^1 \left(\overline{\boldsymbol{\mu}}_2^1 - C\boldsymbol{z}_2^1\right) \tag{2.30}$$

$$\boldsymbol{\mu}_{\boldsymbol{x}(2)}^2 = \overline{\boldsymbol{\mu}}_2^1 + K_1^1 \left(\overline{\boldsymbol{\mu}}_2^1 - C\boldsymbol{z}_2^2\right) \tag{2.31}$$

From this representation it can be seen we are essentially running two Kalman Filters in parallel. The will system pursue both possible trajectories introduced by the measurement update. Although this system is set-up to ensure one trajectory will dominate, see Figure 2.1.

2.1.4 And Beyond ...

We have now established, with some rigour (except for g), that from time steps $t > 2$ a mixture of Gaussians is being propagated forwards in time. After each measurement update the number of mixture components double, for a time step n there are $T(n) = 2^{n-1}$ many mixture components. The system is similar to running exponential number of Kalman Filters in parallel, each pursuing the possible trajectories the measurement updates imply. Naturally, most components will decay after time, but they will still be present in the mixture.

Figure 2.2 shows the exponential explosion of mixture components for n time steps. Each of the leaves is a mixture component with its measured trajectory being the path back to the root. The red path is the dominant mixture component.

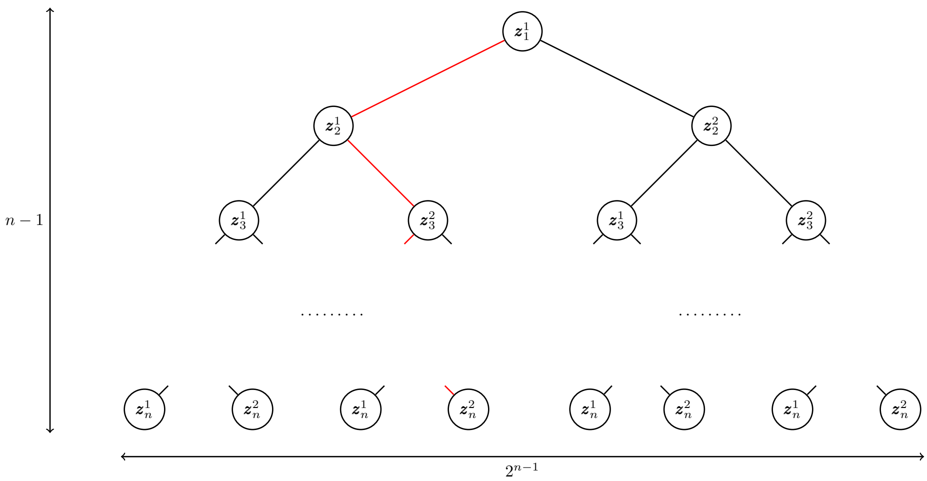
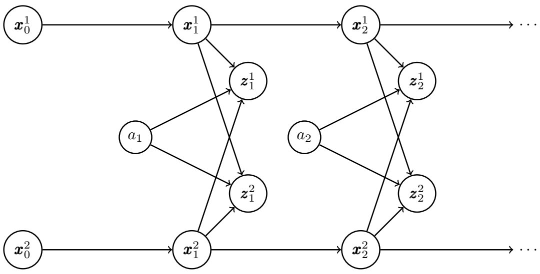


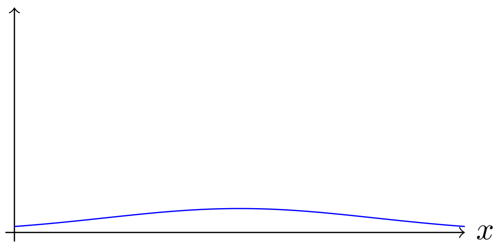
Figure 2.2: A tree enumerating all possible trajectories introduced by measurement updates for n-many time steps. The red path is the dominant trajectory.

2.2 Identifiability

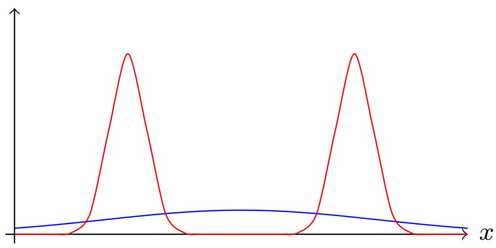
This section is an aside which looks at the previous representation and its identity crisis. Figure 2.2 shows that if we do not initially assign a specific measurement to a target, both will have significant presence in the initial mixture. Since the motion model is the same for both targets the problem becomes symmetrical and $\boldsymbol{x}_t^1$ will be indistinguishable from $\boldsymbol{x}_t^2$. Both mixtures will have two dominant components corresponding to the Kalman Filter's independent estimate of both targets' states, but we will not be able determine which state corresponds to which target. This approach won't have any more mixture components than Figure 2.2, just two dominant components.



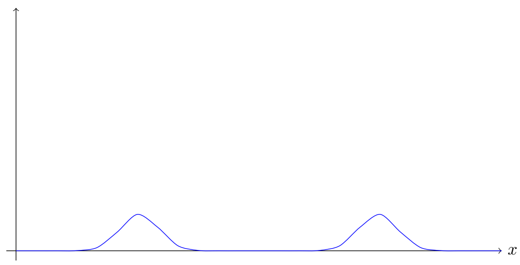
(a) A similar system, which begins with an association problem.



(b) The initial belief of x_1^1, we don't know much about anything.



(c) Before measurement update. Both measurements are likely to been caused by x_1^1.



(d) After the measurement update. Without more specific knowledge about x_1^1's state, both trajectories are likely.

Figure 2.3: An similar system with an identifiability issue. Without specific knowledge of x_0^1 both measurements seem likely and the state of x_t^i will have two dominant components.