The Kalman filter

D.A. Forsyth

University of Illinois at Urbana Champaign

Summary

- Modelling distributions is just maintenance on
  - means, covariances
  - IF
    - prior is normal
    - dynamics are linear + Gaussian noise
    - measurement is linear+Gaussian noise
- which means
  - can obtain the best possible online estimate of state
    - given sequence of observations
    - with a straightforward recipe
- consequences
  - recipe for online procedures (tracking, SLAM)

Linear dynamics and measurement

- State changes as:

Square matrix of full rank

$$\mathbf{x}_i = \mathcal{D}_i \mathbf{x}_{i-1} + \xi$$

This is a normal random variable with zero mean and known covariance

Any matrix whose dimensions are OK

$$\mathbf{y}_i = \mathcal{M}_i \mathbf{x}_i + \zeta$$

- Measurements are:

This is a (different!) normal random variable with zero mean and known covariance

Other notation

Read this as: x_i is normally distributed. The mean is a linear function of x_i-1 and the covariance is known (and can depend on i).

$$\mathbf{x}_i \sim N(\mathcal{D}_i \mathbf{x}_{i-1}; \Sigma_{di})$$

$$\mathbf{y}_i \sim N(\mathcal{M}_i \mathbf{x}_i; \Sigma_{mi})$$

Read this as: y_i is normally distributed. The mean is a linear function of x_i and the covariance is known (and can depend on i)

Example dynamical models

- Drifting points
  - new state = old state + gaussian noise
- Points moving with constant velocity
  - new position=old position + (dt) old velocity + gaussian noise
  - new velocity= old velocity+gaussian noise
- Points moving with constant acceleration
  - new position=old position + (dt) old velocity + gaussian noise
  - new velocity= old velocity+(dt) acceleration + gaussian noise

Example measurement models

- Usual:
  - state=position; measurement=position+gaussian noise
  - state=position and velocity; measurement=position+gaussian noise
    - but we could infer velocity
  - state=position and velocity and acceleration; measurement=position+gaussian noise
    - but we could infer velocity
    - ballistic movement – we know acceleration
- For more complex examples
  - there is a theory of observability
    - can you determine state from observations?

Key point

- $P(\mathbf{y}_i|\mathbf{x}_i)$ is normal.
- If $P(\mathbf{x}_{i-1}|\mathbf{y}_0,...,\mathbf{y}_{i-1})$ is normal
  - then

$$P(\mathbf{x}_i|\mathbf{y}_0,...,\mathbf{y}_{i-1})$$

$$P(\mathbf{x}_i|\mathbf{y}_0,...,\mathbf{y}_i)$$

are both normal

Checking...

- Probability distribution is normal iff it has the form:

$$\log p(\mathbf{x}) = -\frac{1}{2}\left[(\mathbf{x}-\mu)^T \Sigma^{-1} (\mathbf{x}-\mu)\right] + K$$

  - and you can check this for each of the relevant dists.

The Kalman Filter

- Dynamics and measurement

$$\mathbf{x}_i \sim N(\mathcal{D}_i \mathbf{x}_{i-1}, \Sigma_{d_i})$$
$$\mathbf{y}_i \sim N(\mathcal{M}_i \mathbf{x}_i, \Sigma_{m_i})$$

- Notation

$$\text{mean of } P(X_i | y_0, \ldots, y_{i-1}) \text{ as } \bar{X}_i^-$$
$$\text{covar of } P(X_i | y_0, \ldots, y_{i-1}) \text{ as } \Sigma_i^-$$

$$\text{mean of } P(X_i | y_0, \ldots, y_i) \text{ as } \bar{X}_i^+$$
$$\text{covar of } P(X_i | y_0, \ldots, y_i) \text{ as } \Sigma_i^+$$

The steps:

Have:

Mean and covariance of posterior
after i-1'th measurement

Construct:

Mean and covariance of predictive
distribution just before i'th measurement

Measurement arrives:

Now construct:

Mean and covariance of posterior
distribution just before i'th measurement

posterior mean is weighted combo
of prior mean and measurement

posterior covar is weighted combo
of prior covar, measurement
matrix and measurement covar

The steps:

Have:

$$\overline{X}^{+}_{i-1} \qquad \Sigma^{+}_{i-1}$$

Construct:

$$\overline{X}^{-}_{i} = \mathcal{D}_i \overline{X}^{+}_{i-1} \qquad \Sigma^{-}_{i} = \Sigma_{d_i} + \mathcal{D}_i \Sigma^{-}_{i-1} \mathcal{D}_i^{T}$$

Measurement arrives: $\mathbf{y}_i \sim N(\mathcal{M}_i \mathbf{x}_i; \Sigma_{mi})$

Now construct:

$$\bar{X}^{+}_{i} = \bar{X}^{-}_{i} + \mathcal{K}_i \left[\mathbf{y}_i - \mathcal{M}_i \bar{X}^{-}_{i}\right] \qquad \Sigma^{+}_{i} = \left[\mathcal{I} - \mathcal{K}_i \mathcal{M}_i\right] \Sigma^{-}_{i}$$

Where:

$$\mathcal{K}_i = \Sigma^{-}_{i} \mathcal{M}_i^{T} \left[\mathcal{M}_i \Sigma^{-}_{i} \mathcal{M}_i^{T} + \Sigma_{m_i}\right]^{-1}$$

Very simple example

- We have a point vehicle in 2D
    - so steering direction doesn't matter
    - you can accelerate in any dir.
- It is translating
    - we supply a known demand to the accelerator,
        - changing at each time step
    - it sees 2 beacons (which are in its coordinate system)
        - beacon 1 measured in vehicle x but not y
        - beacon 2 measured in vehicle y but not x
- Q:
    - recover filtered estimates of:
        - position, velocity and acceleration in world coords

Dynamical model

- We supply a demand to the accelerator
  - acceleration updates as noise
    - regard demand as a measurement

$$\mathbf{a}_{i+1} = \mathbf{a}_i + \text{noise}$$

  - velocity by integrating acceleration

$$\mathbf{v}_{i+1} = \mathbf{v}_i + \delta t \mathbf{a}_i + \text{noise}$$

  - position by integrating velocity

$$\mathbf{c}_{i+1} = \mathbf{c}_i + \delta t \mathbf{v}_i + \text{noise}$$

Stack the vectors to get:

$$\mathbf{x}_i = \begin{bmatrix} \mathbf{c}_i \\ \mathbf{v}_i \\ \mathbf{a}_i \end{bmatrix}$$

Which gives:

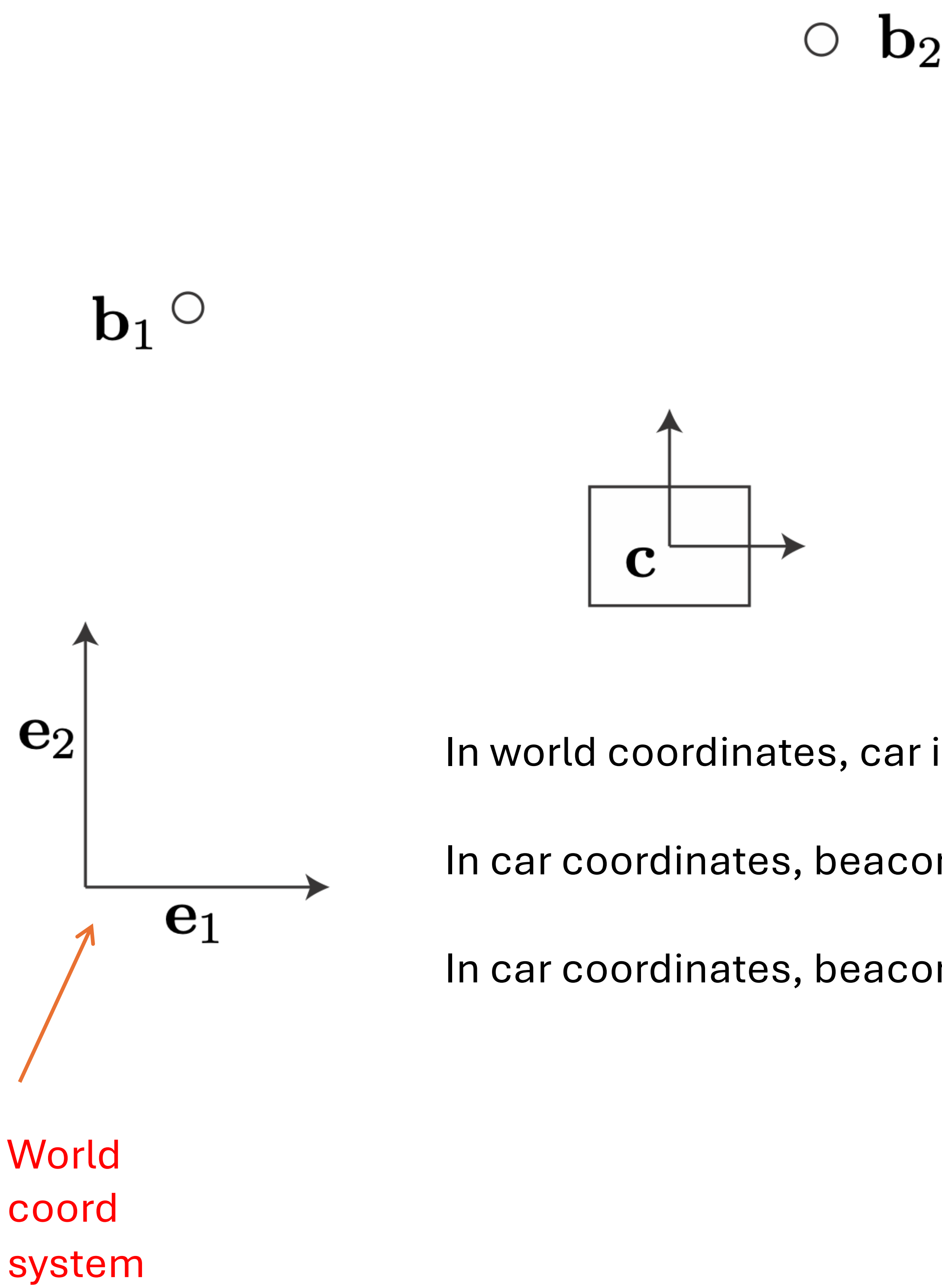
$$\mathbf{x}_{i+1} = \begin{bmatrix} \mathbf{c}_{i+1} \\ \mathbf{v}_{i+1} \\ \mathbf{a}_{i+1} \end{bmatrix} = \begin{bmatrix} \mathcal{I} & \delta t \mathcal{I} & 0 \\ 0 & \mathcal{I} & \delta t \mathcal{I} \\ 0 & 0 & \mathcal{I} \end{bmatrix} \begin{bmatrix} \mathbf{c}_i \\ \mathbf{v}_i \\ \mathbf{a}_i \end{bmatrix} + \xi_i$$

Where:

$$\xi_i \sim N(\mathbf{0}; \Sigma_{d,i})$$

Measurement model

- The acceleration at i should be demand
    - +noise
- Beacons are in car coordinate system)
    - beacon 1 measured in car x but not y
    - beacon 2 measured in car y but not x



In world coordinates, car is at: $\mathbf{c}$

In car coordinates, beacon 1 measurement is: $\mathbf{e}_1^T(\mathbf{b}_1 - \mathbf{c})$

In car coordinates, beacon 2 measurement is: $\mathbf{e}_2^T(\mathbf{b}_2 - \mathbf{c})$

What you measure

The acceleration demand

$$\mathbf{y}_i = \begin{bmatrix} \mathbf{d}_i \\ \mathbf{e}_1^T \mathbf{b}_1 - b_1 \\ \mathbf{e}_2^T \mathbf{b}_2 - b_2 \end{bmatrix} + \text{noise}$$

These are known constants

Measurements from the beacons

Observation model

$$\mathbf{y}_i = \begin{bmatrix} 0 & 0 & \mathcal{I} \\ \mathbf{e}_1^T & 0 & 0 \\ \mathbf{e}_2^T & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{c}_i \\ \mathbf{v}_i \\ \mathbf{a}_i \end{bmatrix} + \text{noise} = \begin{bmatrix} 0 & 0 & \mathcal{I} \\ \mathbf{e}_1^T & 0 & 0 \\ \mathbf{e}_2^T & 0 & 0 \end{bmatrix} \mathbf{x}_i + \zeta_i$$

$$\zeta_i \sim N(0; \Sigma_{m,i})$$

The steps:

Have:

$$\overline{X}_{i-1}^{+} \qquad \Sigma_{i-1}^{+}$$

Construct:

$$\overline{X}_{i}^{-} = \mathcal{D}_i \overline{X}_{i-1}^{+} \qquad \Sigma_{i}^{-} = \Sigma_{d_i} + \mathcal{D}_i \Sigma_{i-1}^{-} \mathcal{D}_i^{T}$$

Measurement arrives: $\mathbf{y}_i \sim N(\mathcal{M}_i \mathbf{x}_i; \Sigma_{mi})$

Now construct:

$$\bar{X}_i^{+} = \bar{X}_i^{-} + \mathcal{K}_i \left[\mathbf{y}_i - \mathcal{M}_i \bar{X}_i^{-}\right] \qquad \Sigma_i^{+} = \left[\mathcal{I} - \mathcal{K}_i \mathcal{M}_i\right] \Sigma_i^{-}$$

Where:

$$\mathcal{K}_i = \Sigma_i^{-} \mathcal{M}_i^{T} \left[\mathcal{M}_i \Sigma_i^{-} \mathcal{M}_i^{T} + \Sigma_{m_i}\right]^{-1}$$

Summary

- Modelling distributions is just maintenance on
  - means, covariances
  - IF
    - prior is normal
    - dynamics are linear + Gaussian noise
    - measurement is linear+Gaussian noise
- which means
  - can obtain the best possible online estimate of state
    - given sequence of observations
    - with a straightforward recipe
- consequences
  - recipe for online procedures (tracking, SLAM)