EKF and EKF SLAM

D.A. Forsyth, UIUC

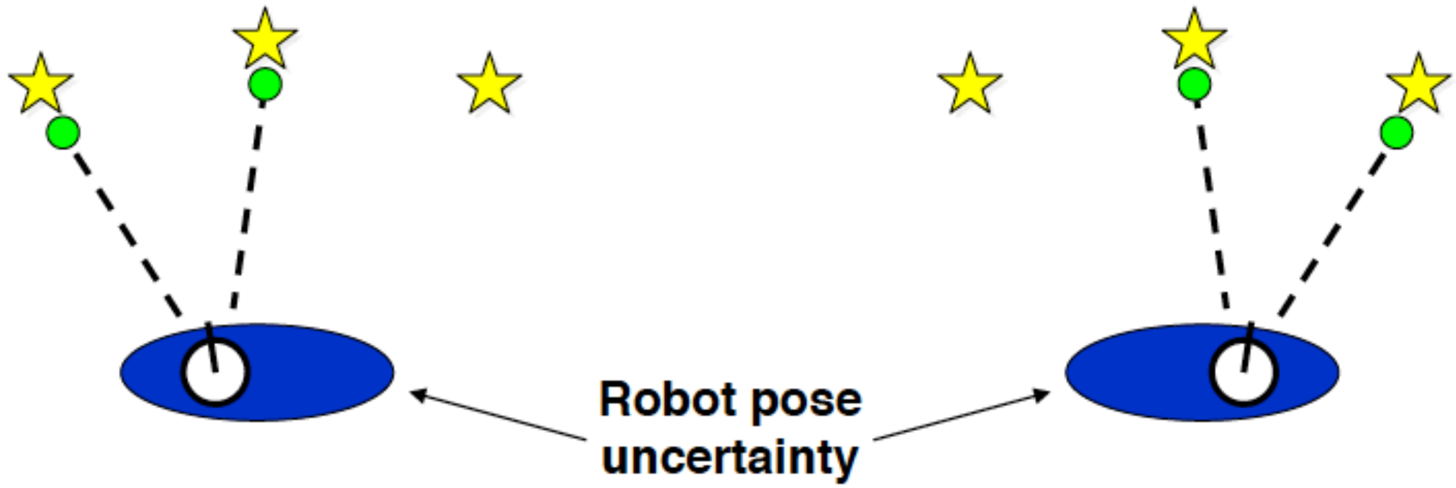
The SLAM Problem

- SLAM stands for simultaneous localization and mapping
- The task of building a map while estimating the pose of the robot relative to this map

- Why is SLAM hard?
  Chicken-or-egg problem:
  - a map is needed to localize the robot and
    a pose estimate is needed to build a map

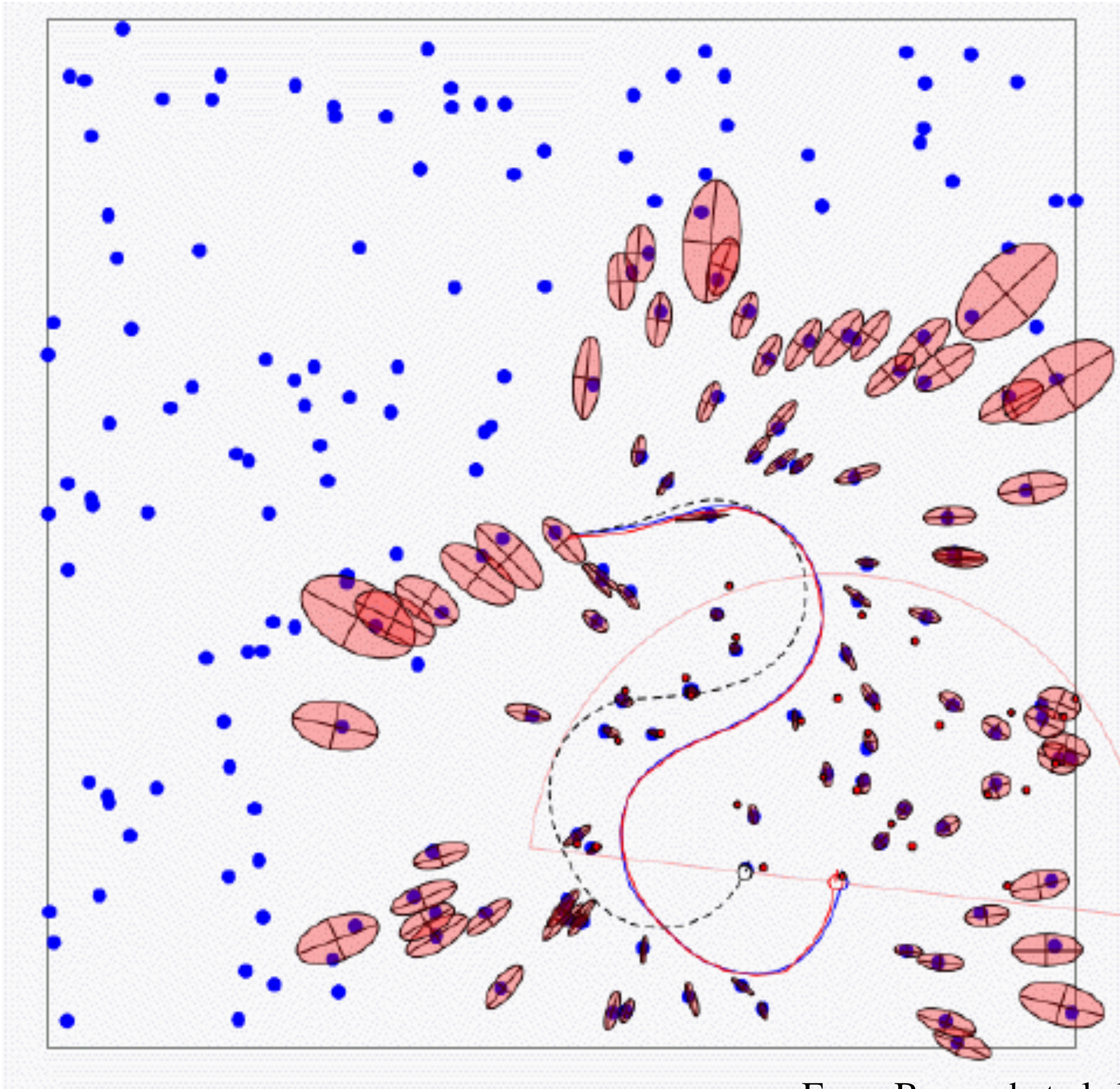
From Burgard et al slides

Why is SLAM a hard problem?



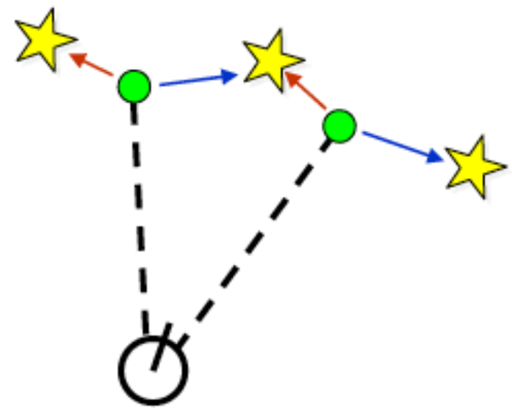
- In the real world, the mapping between observations and landmarks is unknown
- Picking wrong data associations can have catastrophic consequences
- Pose error correlates data associations

From Burgard et al slides



From Burgard et al slides

Data Association Problem



- A data association is an assignment of observations to landmarks
- In general there are more than $\binom{n}{m}$ (n observations, m landmarks) possible associations
- Also called "assignment problem"

From Burgard et al slides

State

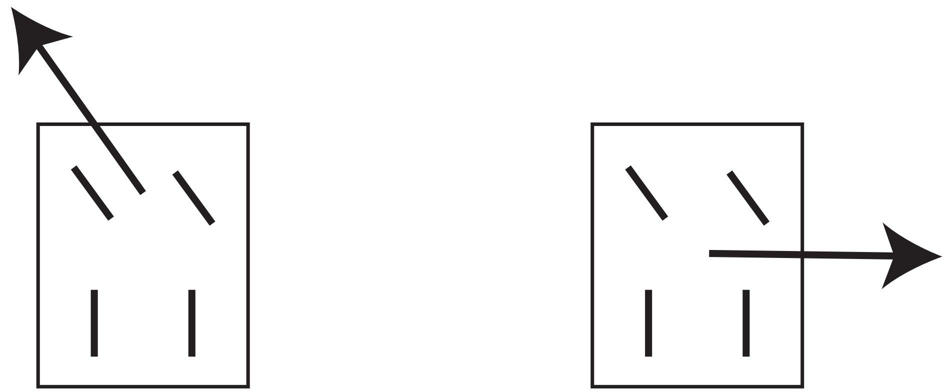
Position and orientation of the robot

$$\mathbf{x} = \begin{bmatrix} \mathcal{R} \\ \mathcal{M} \end{bmatrix} = \begin{bmatrix} \mathcal{R} \\ \mathcal{L}_1 \\ \dots \\ \mathcal{L}_n \end{bmatrix}$$

Landmark 1 position in OCF

All landmark positions in original coordinate frame

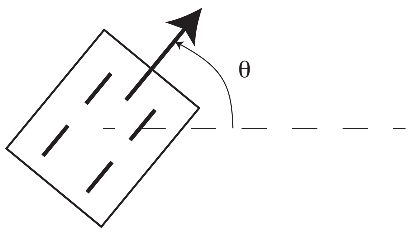
A movement model



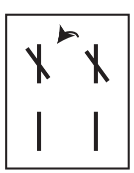
OK　　　　Not OK

Formally: car is non-holonomic

A movement model

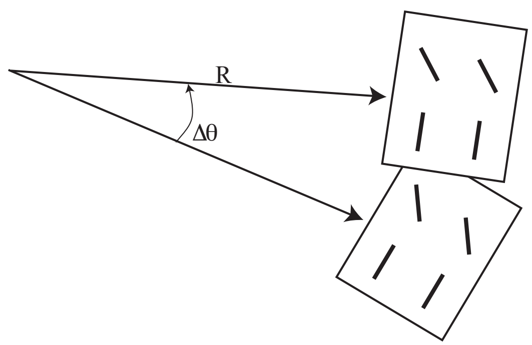


$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \rightarrow \begin{bmatrix} x + v\Delta t cos\theta \\ y + v\Delta t \sin\theta \\ \theta \end{bmatrix}$$



$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ \theta + \Delta\theta \end{bmatrix}$$

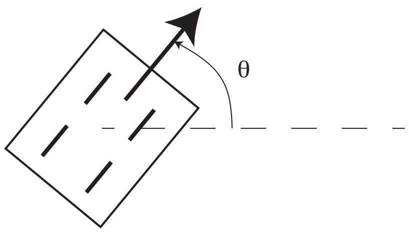
A movement model



THIS ISN'T LINEAR!

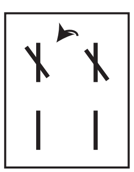
$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \to \begin{bmatrix} x + R(\sin(\theta + \Delta\theta) - \sin\theta) \\ y - R(\cos(\theta + \Delta\theta) - \cos\theta) \\ \theta + \Delta\theta \end{bmatrix}$$

A movement model



$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \rightarrow \begin{bmatrix} x + v\Delta t cos\theta \\ y + v\Delta t \sin\theta \\ \theta \end{bmatrix}$$

These two are limits of previous model (\delta \theta ->0; R->0)



$$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix} \rightarrow \begin{bmatrix} x \\ y \\ \theta + \Delta\theta \end{bmatrix}$$

One kind of measurement model

- Landmark is at:
  - in global coordinate system
- We record distance and heading:
  - measurement

$$\begin{bmatrix} u \\ v \end{bmatrix}$$

$$\begin{bmatrix} d \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{(x-u)^2 + (y-v)^2} \\ \text{atan2}(y-u, x-v) - \theta \end{bmatrix}$$

THIS ISN'T LINEAR!

Another kind of measurement model

- Landmark is at:
  - in global coordinate system
- We record position in vehicle's frame:

$$\left[\begin{array}{c} u \\ v \end{array}\right]$$

$$\left[\begin{array}{c} x_v \\ y_v \end{array}\right] = \mathcal{R}_{-\theta} \left[\begin{array}{c} (u-x) \\ (v-y) \end{array}\right]$$

THIS ISN'T LINEAR!

Linearization and noise

- we have noise $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \Sigma)$
- this means $f(\mathbf{x}+\mathbf{n})$ is a random variable
- Write

$$J_{f,x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_j} & \dots \\ \dots & \frac{\partial f_i}{\partial x_j} & \dots & \dots \\ \dots & \dots & \dots & \frac{\partial f_r}{\partial x_s} \end{bmatrix}$$

- Then

$$f(\mathbf{x}+\mathbf{n}) \approx f(\mathbf{x}) + J_{f,x}\mathbf{n}$$

- So (approximately)

$$f(\mathbf{x}+\mathbf{n}) \sim \mathcal{N}(f(\mathbf{x}), J_{f,x}\Sigma J_{f,x}^T)$$

Linearization and noise

- we have noise $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \Sigma)$

- So (approximately)

$$f(\mathbf{x}, \mathbf{n}) \sim \mathcal{N}(f(\mathbf{x}, \mathbf{0}), J_{f,n} \Sigma J_{f,n}^T)$$

The Kalman filter

Assumption: state update and measurement are linear with normal noise

Dynamic Model:

$$x_i \sim N(\mathcal{D}_i x_{i-1}, \Sigma_{d_i})$$
$$y_i \sim N(\mathcal{M}_i x_i, \Sigma_{m_i})$$

Start Assumptions: $\overline{x}_0^-$ and Σ_0^- are known
Update Equations: Prediction

$$\overline{x}_i^- = \mathcal{D}_i \overline{x}_{i-1}^+$$
$$\Sigma_i^- = \Sigma_{d_i} + \mathcal{D}_i \sigma_{i-1}^+ \mathcal{D}_i$$

Update Equations: Correction

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

Difference between predicted and observed measurement

Algorithm 11.3: The Kalman Filter.

The extended Kalman filter

- What happens if state update, measurement aren't linear?
  - particle filter
  - linearize and approximate (EKF)

$$\mathbf{x}_i = f(\mathbf{x}_{i-1}, \mathbf{n})$$

↑

Noise - normal, mean 0, Cov known

↓

$$\mathbf{y}_i = g(\mathbf{x}_i, \mathbf{n})$$

The extended Kalman filter

- Linearize:

$$\mathbf{x}_i = f(\mathbf{x}_{i-1}, \mathbf{n})$$

$$\mathcal{F}_x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \cdots \\ \cdots & \frac{\partial f_i}{\partial x_j} & \cdots \end{bmatrix}$$

$$\mathcal{F}_n = \begin{bmatrix} \frac{\partial f_1}{\partial n_1} & \cdots & \cdots \\ \cdots & \frac{\partial f_i}{\partial n_j} & \cdots \end{bmatrix}$$

Posterior covariance of x_{i-1}

$$\mathbf{x}_i \sim \mathcal{N}(f(\mathbf{x}_{i-1}, \mathbf{0}), \mathcal{F}_x \Sigma^+_{i-1} \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Noise covariance

Dynamic Model:

$$\boldsymbol{y}_i \quad \sim N(\mathcal{M}_i \boldsymbol{x}_i, \Sigma_{m_i})$$

Start Assumptions: $\overline{\boldsymbol{x}}_0^-$ and Σ_0^- are known

Update Equations: Prediction $\overline{x}_i^-$ Σ_i^-

$$\mathbf{x}_i \sim \mathcal{N}(f(\mathbf{x}_{i-1}, \mathbf{0}), \mathcal{F}_x \Sigma_{i-1}^+ \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Update Equations: Correction

$$\begin{aligned}
\mathcal{K}_i &= \Sigma_i^- \mathcal{M}_i^T \left[\mathcal{M}_i \Sigma_i^- \mathcal{M}_i^T + \Sigma_{m_i}\right]^{-1} \\
\overline{\boldsymbol{x}}_i^+ &= \overline{\boldsymbol{x}}_i^- + \mathcal{K}_i \left[\boldsymbol{y}_i - \mathcal{M}_i \overline{\boldsymbol{x}}_i^-\right] \\
\Sigma_i^+ &= \left[Id - \mathcal{K}_i \mathcal{M}_i\right] \Sigma_i^-
\end{aligned}$$

Algorithm 11.3: The Kalman Filter.

The extended Kalman filter

- Linearize:

$$\mathbf{y}_i = g(\mathbf{x}_i, \mathbf{n})$$

$$\mathcal{G}_x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \cdots \\ \cdots & \frac{\partial f_i}{\partial x_j} & \cdots \end{bmatrix}$$

$$\mathcal{G}_n = \begin{bmatrix} \frac{\partial f_1}{\partial n_1} & \cdots & \cdots \\ \cdots & \frac{\partial f_i}{\partial n_j} & \cdots \end{bmatrix}$$

$$\mathbf{y}_i \sim \mathcal{N}(f(\mathbf{x}_i, \mathbf{0}), \mathcal{G}_x \Sigma_i^- \mathcal{G}_x^T + \mathcal{G}_n \Sigma_{m,i} \mathcal{G}_n^T)$$

Dynamic Model:

$$y_i \sim N(\mathcal{M}_i x_i, \Sigma_{m_i})$$

Start Assumptions: $\overline{x}_0^-$ and Σ_0^- are known

Update Equations: Prediction

$\overline{x}_i^-$ Σ_i^-

$$\mathbf{x}_i \sim \mathcal{N}(f(\mathbf{x}_{i-1}, \mathbf{0}), \mathcal{F}_x \Sigma_{i-1}^+ \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Update Equations: Correction

$$\begin{aligned}
\mathcal{K}_i &= \Sigma_i^- \mathcal{M}_i^T \left[\mathcal{M}_i \Sigma_i^- \mathcal{M}_i^T + \Sigma_{m_i}\right]^{-1} \\
\overline{x}_i^+ &= \overline{x}_i^- + \mathcal{K}_i \left[y_i - \mathcal{M}_i \overline{x}_i^-\right] \\
\Sigma_i^+ &= \left[Id - \mathcal{K}_i \mathcal{M}_i\right] \Sigma_i^-
\end{aligned}$$

This is the inverse of the covariance of y_i

Algorithm 11.3: The Kalman Filter.

Dynamic Model:

$$y_i \quad \sim N(\mathcal{M}_i x_i, \Sigma_{m_i})$$

Start Assumptions: $\overline{x}_0^-$ and Σ_0^- are known

Update Equations: Prediction

$\overline{x}_i^-$ $\quad$ Σ_i^-

$$\mathbf{x}_i \sim \mathcal{N}(f(\mathbf{x}_{i-1}, \mathbf{0}), \mathcal{F}_x \Sigma_{i-1}^+ \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Update Equations: Correction

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

$$\overline{x}_i^+ \quad = \overline{x}_i^- + \mathcal{K}_i \left[y_i - \mathcal{M}_i \overline{x}_i^-\right]$$

$$\Sigma_i^+ \quad = [Id - \mathcal{K}_i \mathcal{M}_i] \Sigma_i^-$$

Difference between predicted and true measurement

Algorithm 11.3: The Kalman Filter.

Dynamic Model:

$$y_i \sim N(\mathcal{M}_i x_i, \Sigma_{m_i})$$

Start Assumptions: $\overline{x}_0^-$ and Σ_0^- are known

Update Equations: Prediction

$\overline{x}_i^-$ Σ_i^-

$$\mathbf{x}_i \sim \mathcal{N}(f(\mathbf{x}_{i-1}, \mathbf{0}), \mathcal{F}_x \Sigma_{i-1}^+ \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Update Equations: Correction

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

$$\overline{x}_i^+ = \overline{x}_i^- + \mathcal{K}_i \left[y_i - \mathcal{M}_i \overline{x}_i^-\right]$$

$$\Sigma_i^+ = [Id - \mathcal{K}_i \mathcal{M}_i] \Sigma_i^-$$

Linear measurement model

Algorithm 11.3: The Kalman Filter.

Dynamic Model:

$$\mathbf{x}_i = f(\mathbf{x}_{i-1}, \mathbf{n})$$
$$\mathbf{y}_i = g(\mathbf{x}_i, \mathbf{n})$$

Start Assumptions: $\overline{\boldsymbol{x}}_0^-$ and Σ_0^- are known

Update Equations: Prediction $\overline{x}_i^-$ Σ_i^-

$$\mathbf{x}_i \sim \mathcal{N}(f(\mathbf{x}_{i-1}, \mathbf{0}), \mathcal{F}_x \Sigma_{i-1}^+ \mathcal{F}_x^T + \mathcal{F}_n \Sigma_{n,i} \mathcal{F}_n^T$$

Update Equations: Correction

$$\begin{aligned} \mathcal{K}_i &= \Sigma_i^- \mathcal{M}_i^T \left[\mathcal{G}_x \Sigma_i^- \mathcal{G}_x^T + \mathcal{G}_n \Sigma_{m,i} \mathcal{G}_n^T\right]^{-1} \\ \overline{\boldsymbol{x}}_i^+ &= \overline{\boldsymbol{x}}_i^- + \mathcal{K}_i \left[\mathbf{y}_i - g(\mathbf{x}_i^-, \mathbf{0})\right] \\ \Sigma_i^+ &= [Id - \mathcal{K}_i \mathcal{G}_x] \Sigma_i^- \end{aligned}$$

Algorithm 11.3: The Kalman Filter.

In principle, now easy

- BUT
  - F_x is much simpler than it might look
    - the landmarks do not move!
  - F_n ditto
    - there is no noise in the landmark updates

N=Number of landmarks

$$\mathcal{F}_x = \begin{array}{cc} \underline{\;3\;} & \underline{\;N\;} \end{array}$$

$$\mathcal{F}_x = \begin{bmatrix} \frac{\partial f_{\mathcal{R}}}{\partial \mathcal{R}} & 0 \\ 0 & \mathcal{I} \end{bmatrix}$$

$$\mathcal{F}_n = \begin{bmatrix} \frac{\partial f_{\mathcal{R}}}{\partial \mathbf{n}} & 0 \\ 0 & 0 \end{bmatrix}$$

More simplifications

- BUT
  - G_x is much simpler than it might look
    - each set of measurements affected by only one landmark!

N N=Number of landmarks

$$\mathcal{G}_x = \begin{bmatrix} \overset{3}{\frac{\partial g_{\mathcal{R}}}{\partial \mathcal{R}}} & \overset{3}{\frac{\partial g_{\mathcal{L}_1}}{\partial \mathcal{L}_1}} & 0 & 0 & 0 & 0 \\ \frac{\partial g_{\mathcal{R}}}{\partial \mathcal{R}} & 0 & \frac{\partial g_{\mathcal{L}_2}}{\partial \mathcal{L}_2} & 0 & 0 & 0 \\ \cdots & & & & & \\ \frac{\partial g_{\mathcal{R}}}{\partial \mathcal{R}} & \cdots & \cdots & \frac{\partial g_{\mathcal{L}_i}}{\partial \mathcal{L}_i} & \cdots & \end{bmatrix}$$

More simplifications

- BUT
  - G_n is usually much simpler than it might look
    - noise is usually additive normal noise

Block diagonal

$$\mathcal{G}_n^T \Sigma_{n,i} \mathcal{G}_n \rightarrow \Sigma_{n,i}$$

Landmarks

- Which measurement comes from which landmark?
  - data association -
    - for the moment, assume
      - we use a bipartite graph matcher
      - or draw independent samples from posterior on landmark
        - given measurement
    - ideally, we'd average over all matchings - put that off

Landmarks

- No measurement from a landmark?
  - structure of EKF means you can process landmarks one by one
  - don't update that landmark
- New landmark?
  - full observation (eg range+bearing, lidar)
  - partial observation (eg bearing, vision)

Full observation

- Must make estimates of
  - landmark mean state
    - invert the observation of the landmark
  - landmark covariance
    - with itself
    - with others
    - use jacobians of inverted observation

Range and bearing

Landmark position

Observation → $$\begin{bmatrix} d \\ \phi \end{bmatrix} = \begin{bmatrix} \sqrt{(x-u)^2 + (y-v)^2} \\ \operatorname{atan2}(y-u, x-v) - \theta \end{bmatrix}$$

Vehicle state

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x + d\sin(\phi+\theta) \\ y + d\cos(\phi+\theta) \end{bmatrix} = h(\overline{\mathbf{x}}_i^-, \mathbf{y}_i) = h(\overline{\mathbf{x}}_i^-, \mathbf{l})$$

Here use the current estimate of vehicle state

These are measurements of new landmark ONLY

Range and bearing

- but the measurement may be affected by noise
  - additive noise, normal, zero mean, covar $\sum$
- So I should have written

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} x + (d+\xi)\sin(\phi+\zeta+\theta) \\ y + (d+\xi)\cos(\phi+\zeta+\theta) \end{bmatrix} = h(\overline{\mathbf{x}}_i^-, \mathbf{y}_i, \xi, \zeta) = h(\overline{\mathbf{x}}_i^-, \mathbf{l}, \mathbf{n})$$

- And I need to do some surgery
  - on the state vector
  - and on the covariance matrix

Range and bearing - state vector surgery

- Because the noise has zero mean

$$\mathbf{x}_i^+ \rightarrow \begin{bmatrix} \mathbf{x}_i^+ \\ u \\ v \end{bmatrix}$$

Range and bearing - covariance surgery

$$\left[\begin{array}{c} u \\ v \end{array}\right] = \left[\begin{array}{c} x + (d+\xi)\sin(\phi+\zeta+\theta) \\ y + (d+\xi)\cos(\phi+\zeta+\theta) \end{array}\right] = h(\overline{\mathbf{x}}_i^-, \mathbf{y}_i, \xi, \zeta) = h(\overline{\mathbf{x}}_i^-, \mathbf{l}, \mathbf{n})$$

- So

Covariance of vehicle state with itself ↓

$$\mathcal{H}_x \Sigma_{i,xx}^{+} \mathcal{H}_x^T + \mathcal{H}_n \Sigma \mathcal{H}_n^T$$

← Covariance of landmark with itself

↑ Jacobian of landmark position wrt vehicle state

- and

↓

$$\mathcal{H}_x \Sigma_{i,\mathcal{RM}}^{+}$$

← Covariance of landmark with everything else

↑ i'th posterior covariance of location with all other landmarks

Range and bearing - covariance surgery

$$\Sigma_i^+ \rightarrow \left[\begin{array}{cc} \Sigma_i^+ & \left(\mathcal{H}_x \Sigma_{i,\mathcal{RM}}^+\right)^T \\ \left(\mathcal{H}_x \Sigma_{i,\mathcal{RM}}^+\right) & \mathcal{H}_x \Sigma_{i,xx}^+ \mathcal{H}_x^T + \mathcal{H}_n \Sigma \mathcal{H}_n^T \end{array} \right]$$

Bearing only (sketch)

- Cannot determine landmark in 2D from measurement
  - it's on a line!
  - you must come up with a prior
    - after that, it's easy
      - find mean posterior location, covariance
      - plug in
  - Big Issue
    - True prior should have infinite covariance
      - can't work with that
      - so linearization may fail