Extended Kalman Filter for SLAM

Sami Osborn

1 Mathematical Foundations of EKF-SLAM

Simultaneous Localisation and Mapping (SLAM) seeks to estimate both a robot's pose and the positions of environmental landmarks. The Extended Kalman Filter (EKF) provides a probabilistic framework for doing this in real time, by maintaining a joint Gaussian distribution over all unknowns and updating it recursively as new observations arrive.

At its core, the EKF combines:

- A **motion model** describing how the robot's state evolves.
- A **measurement model** describing how landmarks appear in the camera or sensor frame.

The following builds up the mathematics of EKF-SLAM from its probabilistic foundations.

2 Probabilistic Model

Let the robot's pose at time t be $x_t \in SE(3)$, and the map consist of N static landmarks $m_i \in \mathbb{R}^3$, collected in a global state vector:

$$X_t = \begin{bmatrix} x_t \\ m_1 \\ m_2 \\ \vdots \\ m_N \end{bmatrix}$$

The system dynamics and measurement processes are modelled as:

$$x_t = f(x_{t-1}, u_t) + w_t, \qquad w_t \sim \mathcal{N}(0, Q_t)$$

$$z_t = h(x_t, M) + v_t, \qquad v_t \sim \mathcal{N}(0, R_t)$$

where u_t is the control input, w_t the process noise, and v_t the measurement noise.

The goal is to recursively estimate the posterior:

$$p(X_t \mid z_{1:t}, u_{1:t})$$

which expresses our belief about the joint robot–map state given all past observations and controls.

3 Gaussian Assumption and Linearisation

The EKF assumes that the posterior is approximately Gaussian:

$$p(X_t \mid z_{1:t}) \approx \mathcal{N}(\hat{X}_t, P_t)$$

where:

$$\hat{X}_t = \mathbb{E}[X_t \mid z_{1:t}], \quad P_t = \mathrm{Var}[X_t \mid z_{1:t}]$$

Because both $f(\cdot)$ and $h(\cdot)$ are nonlinear, we linearise them about the current estimate using first-order Taylor expansion.

For the motion model:

$$f(x_{t-1}, u_t) \approx f(\hat{x}_{t-1}, u_t) + F_t(x_{t-1} - \hat{x}_{t-1})$$

where $F_t = \frac{\partial f}{\partial x}\big|_{\hat{x}_{t-1},u_t}$ is the Jacobian of the transition function.

For the observation model:

$$h(x_t, M) \approx h(\hat{x}_t, \hat{M}) + H_t(X_t - \hat{X}_t)$$

where $H_t = \frac{\partial h}{\partial X}\big|_{\hat{X}_t}$ is the measurement Jacobian.

Jacobian Matrices

The Jacobian is a matrix of first-order partial derivatives that describes how a vector-valued function changes with respect to its input variables. For a nonlinear function

$$f : \mathbb{R}^n \to \mathbb{R}^m, \quad y = f(x)$$

the Jacobian $J_f(x)$ is defined as:

$$J_f(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

Each element $\frac{\partial f_i}{\partial x_j}$ expresses how the i-th output changes with respect to the j-th input. In the EKF, this matrix locally linearises nonlinear functions so that Gaussian uncertainty can be propagated through them.

4 Recursive Estimation

The EKF proceeds in two stages: prediction and update.

Prediction Step

From the system model:

$$\hat{X}_t^- = f(\hat{X}_{t-1}, u_t)$$
$$P_t^- = F_t P_{t-1} F_t^\top + Q_t$$

where $(\hat{X}_t^-, P_t^-)$ denote the predicted mean and covariance before the new observation.

Measurement Update

Given a new observation z_t, the predicted measurement is:

$$\hat{z}_t = h(\hat{X}_t^-)$$

where h is the measurement function which maps the state to sensor readings.

The innovation (residual) is:

$$r_t = z_t - \hat{z}_t$$

where z_t is the actual measurement received from the sensor.

The Kalman gain weights how much we trust the new measurement:

$$K_t = P_t^- H_t^\top (H_t P_t^- H_t^\top + R_t)^{-1}$$

where H_t is the Jacobian of the measurement function h evaluated at $\hat{X}_t^-$, and R_t is the measurement noise covariance.

The posterior mean and covariance become:

$$\hat{X}_t = \hat{X}_t^- + K_t r_t$$

$$P_t = (I - K_t H_t) P_t^-$$

The gain K_t automatically balances prediction confidence versus measurement reliability: large R_t (noisy sensors) reduces K_t, whereas large P_t^- (uncertain prediction) increases it.

5 EKF-SLAM State Structure

In SLAM, the state vector couples the robot pose and all landmarks. Hence P_t is a large joint covariance matrix:

$$P_t = \begin{bmatrix} P_{xx} & P_{xm} \\ P_{mx} & P_{mm} \end{bmatrix}$$

where:

- P_{xx}: uncertainty of robot pose.
- P_{mm}: uncertainty of landmark positions.
- P_{xm}: cross-correlation between pose and landmarks.

When the robot moves, the motion model affects only P_{xx} and its correlations with landmarks. When it observes features, the measurement model couples the corresponding landmark states with the current pose through the Jacobian H_t.

This joint structure allows information from one observation to reduce uncertainty across all correlated landmarks.

6 3D Extension via Pinhole Projection Theorem

For visual SLAM, each measurement is the image-plane projection of a 3D world landmark p_W. Transforming from world to camera coordinates:

$$p_C = R_{CW}(p_W - t_{CW})$$

Under the pinhole projection model:

$$z_{\text{pred}} = h(p_C) = \begin{bmatrix} f_x \frac{X_c}{Z_c} + c_x \\ f_y \frac{Y_c}{Z_c} + c_y \end{bmatrix}$$

The Jacobian of this projection with respect to the 3D point is:

$$\frac{\partial(u, v)}{\partial(X_c, Y_c, Z_c)} = \begin{bmatrix} \frac{f_x}{Z_c} & 0 & -f_x \frac{X_c}{Z_c^2} \\ 0 & \frac{f_y}{Z_c} & -f_y \frac{Y_c}{Z_c^2} \end{bmatrix}$$

This 2×3 matrix captures how small displacements in 3D camera space affect pixel coordinates. Its $1/Z_c$ scaling shows that uncertainty grows with depth, so nearby features provide stronger geometric constraints than distant ones.

The full measurement Jacobian used in the EKF update combines this with derivatives of the camera pose:

$$H_t = \frac{\partial h}{\partial X_t} = \frac{\partial h}{\partial p_C} \frac{\partial p_C}{\partial X_t}$$

7 Geometric and Statistical Interpretation

The EKF's mathematics tightly links projective geometry and probabilistic inference:

- The **camera model** defines how 3D structure is mapped into 2D pixel space.
- The **Jacobian** translates infinitesimal 3D motion into image changes, determining how uncertainty propagates.
- The **Kalman gain** enforces optimal fusion under Gaussian assumptions, adjusting belief in each measurement based on its projected information content.

Thus, EKF-SLAM can be viewed as performing recursive conditioning of a joint Gaussian over robot and landmark states, constrained by projective geometry.