The Kalman filter and LMMSE estimators

Sensor fusion & nonlinear filtering

Lars Hammarstrand

THE KALMAN FILTER IS AN LMMSE ESTIMATOR

- The Kalman filter computes

$$\hat{\mathbf{x}}_{k|k-1} = \mathbb{E}\{\mathbf{x}_k | \mathbf{y}_{1:k-1}\}$$
$$\hat{\mathbf{x}}_{k|k} = \mathbb{E}\{\mathbf{x}_k | \mathbf{y}_{1:k}\}$$

$$\hat{x}_{k|k}(y_{1:k}) = \hat{x}_{k|k-1}(y_{1:k-1}) + K_k(y_k - \hat{y}_{k|k-1}(y_{1:k-1}))$$

for linear and Gaussian models.

- Note:
  1. The Kalman filter is a linear function of $\mathbf{y}_{1:k}$.
  2. $\hat{\mathbf{x}}_{k|k}$ is the minimum mean square error (MMSE) estimator.

$\Rightarrow$ The Kalman filter is the linear minimum mean square error (LMMSE) estimator!

LMMSE ESTIMATION

LMMSE objective (static example)

- Find $\mathbf{A}$ and $\mathbf{b}$ such that $\hat{\mathbf{x}} = \mathbf{A}\mathbf{y} + \mathbf{b}$ yields the smallest possible MSE, $\mathbb{E}\{(\mathbf{x} - \hat{\mathbf{x}})^T(\mathbf{x} - \hat{\mathbf{x}})\}$.

- Finding optimum:
  Setting derivatives of MSE with respect to $\mathbf{b}$ and $\mathbf{A}$ to $\mathbf{0}$ yields

$$\mathbf{b} = \bar{\mathbf{x}} - \mathbf{A}\bar{\mathbf{y}}$$

$$\mathbf{A} = \mathbf{P}_{xy}\mathbf{P}_{yy}^{-1}.$$

$$\Rightarrow \hat{x} = Ay + \bar{x} - A\bar{y} = \bar{x} + A(y - \bar{y})$$

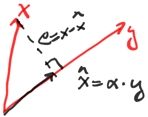
$$= \bar{x} + P_{xx} \cdot P_{yy}^{-1}(y - \bar{y})$$

- Orthogonality principle:
  select $\mathbf{A}$ such that $\mathbb{E}\{(\mathbf{x} - \hat{\mathbf{x}})\mathbf{y}^T\} = \mathbf{0}$.

$$\Leftrightarrow P_{xy} - A \cdot P_{yy} = 0$$

$$\Leftrightarrow A = P_{xy} \cdot P_{yy}^{-1}$$

$E\{xy\}$: inner product $\Rightarrow E\{xy\} = 0 \Leftrightarrow x \perp y$



$x - \hat{x} \perp y \Leftrightarrow E\{(x - \hat{x})y\} = 0$

SELF ASSESSMENT

What is different in LMMSE estimation compared to MMSE estimation.

- In LMMSE estimation we restrict the estimator to be a linear (or at least affine) function of data (measurements).
- In LMMSE the noise cannot be Gaussian.
- In MMSE estimation we normally compute a posterior distribution conditioned on data.

Check all that apply.

SEQUENTIAL LMMSE IN THE DYNAMIC CASE

LMMSE objective

- Sequentially find $\{\mathbf{L}_{k|k-1}, \mathbf{b}_{k|k-1}\}$ and $\{\mathbf{L}_{k|k}, \mathbf{b}_{k|k}\}$ such that

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{L}_{k|k-1}\underline{\mathbf{y}_{1:k-1}} + \mathbf{b}_{k|k-1}$$
$$\hat{\mathbf{x}}_{k|k} = \mathbf{L}_{k|k}\underline{\mathbf{y}_{1:k}} + \mathbf{b}_{k|k}$$

  minimize the MSE, $\mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T(\cdot)\}$ and $\mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T(\cdot)\}$.

Note:

- Find a linear mapping based on all the data up to the relevant time.
- We generalise and allow us to consider affine functions of data.

LMMSE FOR LINEAR STATE SPACE MODELS

Linear state space model with additive (non-Gaussian) noise

- Consider state space model

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{q}_{k-1},$$
$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{r}_k$$

  where $\mathbf{x}_0$, $\mathbf{q}_{k-1}$ and $\mathbf{r}_k$ are independed random variables with known mean and covariances.

Key results (Additive non-Gaussian noise)

- The Kalman filter gives LMMSE estimates, $\hat{\mathbf{x}}_{k|k-1}$ and $\hat{\mathbf{x}}_{k|k}$, with the correct error covariances

$$\mathbf{P}_{k|k-1} = \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\cdot)^T\}$$
$$\mathbf{P}_{k|k} = \mathbb{E}\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\cdot)^T\}$$

PROOF OUTLINE

- **Assumption:** $\begin{cases} \mathbb{E}\left[(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1})\mathbf{y}_{1:k-1}^T\right] = 0 \\ \mathbb{E}\left[(\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1|k-1})(\cdot)^T\right] = \mathbf{P}_{k-1|k-1} \end{cases}$

Prediction	Update
$\mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k\|k-1})\mathbf{y}_{1:k-1}^T\right] = 0$ $\mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k\|k-1})(\cdot)^T\right] = \mathbf{P}_{k\|k-1}$	$\mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k\|k})\mathbf{y}_{1:k-1}^T\right] = \mathbf{0}$ $\mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k\|k})\mathbf{y}_k^T\right] = \mathbf{0}$ $\mathbb{E}\left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k\|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k\|k})^T\right] = \mathbf{P}_{k\|k}$

- **DIY:** Make the proof for the scalar case where x_0, q_{k-1}, r_k are zero mean.

SELF ASSESSMENT

Fact:
For linear state space models with additive noise, the Kalman filter computes the LMMSE estimate recursively, also when the noise is not Gaussian.

Statement for you to verify or reject:
However, the Kalman filter is merely the best linear estimator among all *recursive* algorithms and we can sometime do better if we consider all measurements at the same time.

- True.
- False