卡尔曼滤波器

卡尔曼滤波器是贝叶斯滤波器在**线性高斯系统**下的一个特例。

假设线性高斯系统下运动模型和观测模型如下：

$$\begin{aligned}&\text{运动方程：}\boldsymbol{x}_{k}=\boldsymbol{A}_{k-1} \boldsymbol{x}_{k-1}+\boldsymbol {v}_{k}+\boldsymbol{w}_{k}, \quad k=1, \cdots, K \\&\text{观测方程：}\boldsymbol{y}_{k}=\boldsymbol{C}_{k} x_{k}+\boldsymbol{n}_{k}, \quad k=0, \cdots, K\end{aligned}$$

其中：

- $\boldsymbol{x}_{k} \in \mathbb{R}^{N}$ 表示系统状态
- $\boldsymbol{v}_{k} \in \mathbb{R}^{N}$ 表示输入
- $\boldsymbol{w}_{k} \in \mathbb{R}^{N}$ 表示过程噪声，$\boldsymbol{w}_{k} \sim N(\mathbf{0}, \boldsymbol{Q}_{k})$
- $\boldsymbol{y}_{k} \in \mathbb{R}^{M}$ 表示测量
- $\boldsymbol{n}_{k} \in \mathbb{R}^{M}$ 表示测量噪声，$\boldsymbol{n}_{k} \sim N(\mathbf{0}, \boldsymbol{R}_{k})$
- k 为时间下标，最大值为 K

用 $\check{(\cdot)}$ 表示先验，用 $\hat{(\cdot)}$ 表示后验，卡尔曼滤波器则可以表示为：

预测

$$\check{\boldsymbol{x}}_{k}=\boldsymbol{A}_{k-1} \hat{\boldsymbol{x}}_{k-1}+\boldsymbol{v}_{k}$$

$$\check{\boldsymbol{P}}_{k}=\boldsymbol{A}_{k-1} \hat{\boldsymbol{P}}_{k-1} \boldsymbol{A}_{k-1}^{\mathrm{T}}+\boldsymbol{Q}_{k}$$

卡尔曼增益

$$\quad \boldsymbol{K}_{k}=\check{\boldsymbol{P}}_{k} \boldsymbol{C}_{k}^{\mathrm{T}}\left(\boldsymbol{C}_{k} \check{\boldsymbol{P}}_{k} \boldsymbol{C}_{k}^{\mathrm{T}}+\boldsymbol{R}_{k}\right)^{-1}$$

更新

$$\hat{\boldsymbol{x}}_{k}=\check{\boldsymbol{x}}_{k}+\boldsymbol{K}_{k} \underbrace{\left(\boldsymbol{y}_{k}-\boldsymbol{C}_{k} \check{\boldsymbol{x}}_{k}\right)}_{\text {更新量 }}$$

$$\hat{\boldsymbol{P}}_{k}=\left(\mathbf{1}-\boldsymbol{K}_{k} \boldsymbol{C}_{k}\right) \check{\boldsymbol{P}}_{k}$$

通过贝叶斯推断推导

预测

由于是线性高斯系统，因此直接将 $k-1$ 时刻的后验分布通过线性运动模型传递，即可得到 k 时刻的先验：

$$\begin{aligned}\check{\boldsymbol{x}}_{k} &=E\left[\boldsymbol{x}_{k}\right]\\&=E\left[\boldsymbol{A}_{k-1} \boldsymbol{x}_{k-1}+\boldsymbol{v}_{k}+\boldsymbol{w}_{k}\right] \\&=\boldsymbol{A}_{k-1} \underbrace{E\left[\boldsymbol{x}_{k-1}\right]}_{\hat{\boldsymbol{x}}_{k-1}}+\boldsymbol{v}_{k}+\underbrace{E\left[\boldsymbol{w}_{k}\right]}_{0}\\&=\boldsymbol{A}_{k-1} \hat{\boldsymbol{x}}_{k-1}+\boldsymbol{v}_{k}\end{aligned}$$

$$\begin{aligned}\check{\boldsymbol{P}}_{k}=& E\left[\left(\boldsymbol{x}_{k}-E\left[\boldsymbol{x}_{k}\right]\right)\left(\boldsymbol{x}_{k}-E\left[\boldsymbol{x}_{k}\right]\right)^{\mathrm{T}}\right] \\=& E\left[\left(\boldsymbol{A}_{k-1} \boldsymbol{x}_{k-1}+\boldsymbol{v}_{k}+\boldsymbol{w}_{k}-\boldsymbol{A}_{k-1} \hat{\boldsymbol{x}}_{k-1}-\boldsymbol{v}_{k}\right)\right.\\&\left.\times\left(\boldsymbol{A}_{k-1} \boldsymbol{x}_{k-1}+\boldsymbol{v}_{k}+\boldsymbol{w}_{k}-\boldsymbol{A}_{k-1} \hat{\boldsymbol{x}}_{k-1}-\boldsymbol{v}_{k}\right)^{\mathrm{T}}\right] \\=& \boldsymbol{A}_{k-1} \underbrace{E\left[\left(\boldsymbol{x}_{k-1}-\hat{\boldsymbol{x}}_{k-1}\right)\left(\boldsymbol{x}_{k-1}-\hat{\boldsymbol{x}}_{k-1}\right)^{\mathrm{T}}\right]}_{\hat{\boldsymbol{P}}_{k-1}} \boldsymbol{A}_{k-1}^{\mathrm{T}}+\underbrace{E\left[\boldsymbol{w}_{k} \boldsymbol{w}_{k}^{\mathrm{T}}\right]}_{\boldsymbol{Q}_{k}} \\=& \boldsymbol{A}_{k-1} \hat{\boldsymbol{P}}_{k-1} \boldsymbol{A}_{k-1}^{\mathrm{T}}+\boldsymbol{Q}_{k}\end{aligned}$$

更新

对于更新部分，我们将状态与 k 时刻的测量写成联合高斯分布的形式：

$$\begin{aligned}p\left(\boldsymbol{x}_{k}, \boldsymbol{y}_{k} \mid \check{\boldsymbol{x}}_{0}, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) &=\mathcal{N}\left(\left[\begin{array}{c}\boldsymbol{\mu}_{x} \\\boldsymbol{\mu}_{y}\end{array}\right],\left[\begin{array}{cc}\boldsymbol{\Sigma}_{x x} & \boldsymbol{\Sigma}_{x y} \\\boldsymbol{\Sigma}_{y x} & \boldsymbol{\Sigma}_{y y}\end{array}\right]\right) \\&=\mathcal{N}\left(\left[\begin{array}{c}\check{\boldsymbol{x}}_{k} \\\boldsymbol{C}_{k} \check{\boldsymbol{x}}_{k}\end{array}\right],\left[\begin{array}{cc}\check{\boldsymbol{P}}_{k} & \check{\boldsymbol{P}}_{k} \boldsymbol{C}_{k}^{\mathrm{T}} \\\boldsymbol{C}_{k} \check{\boldsymbol{P}}_{k} & \boldsymbol{C}_{k} \check{\boldsymbol{P}}_{k} \boldsymbol{C}_{k}^{\mathrm{T}}+\boldsymbol{R}_{k}\end{array}\right]\right)\end{aligned}$$

由高斯推断可以直接得到 k 时刻的条件分布（即后验概率）：

$$p\left(\boldsymbol{x}_{k} \mid \check{\boldsymbol{x}}_{0}, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k}\right)=\mathcal{N}(\underbrace{\boldsymbol{\mu}_{x}+\boldsymbol{\Sigma}_{x y} \boldsymbol{\Sigma}_{y y}^{-1}\left(\boldsymbol{y}_{k}-\boldsymbol{\mu}_{y}\right)}_{\hat{\boldsymbol{x}}_{k}}, \underbrace{\boldsymbol{\Sigma}_{x x}-\boldsymbol{\Sigma}_{x y} \boldsymbol{\Sigma}_{y y}^{-1} \boldsymbol{\Sigma}_{y x}}_{\hat{\boldsymbol{P}}_{k}})$$

令 $\boldsymbol{K}_{k}=\boldsymbol{\Sigma}_{x y} \boldsymbol{\Sigma}_{y y}^{-1}=\check{\boldsymbol{P}}_{k} \boldsymbol{C}_{k}^{\mathrm{T}}\left(\boldsymbol{C}_{k} \check{\boldsymbol{P}}_{k} \boldsymbol{C}_{k}^{\mathrm{T}}+\boldsymbol{R}_{k}\right)^{-1}$（卡尔曼增益），则有：

$$\hat{\boldsymbol{x}}_{k}=\check{\boldsymbol{x}}_{k}+\boldsymbol{K}_{k} \left(\boldsymbol{y}_{k}-\boldsymbol{C}_{k} \check{\boldsymbol{x}}_{k}\right)$$

$$\hat{\boldsymbol{P}}_{k}=\left(\mathbf{1}-\boldsymbol{K}_{k} \boldsymbol{C}_{k}\right) \check{\boldsymbol{P}}_{k}$$

即证。