卡尔曼滤波

董开明 王秀程
海洋 柴智 曹思齐

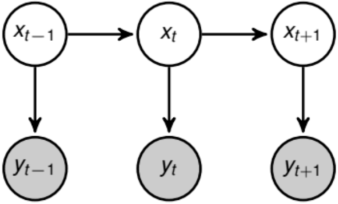
2019 年 11 月 21 日

卡尔曼滤波的引入及主要思想

在定位问题中，由于实际的位置受到噪声的干扰，会不断抖动，即可将其实际位置描述为随机变量，为了对确定其实际位置，我们需要用观测值去估计这个随机变量的均值，根据均方误差最小原则，我们得出其实际位置的最优估计就是这个随机变量的均值。为了得到这样一个均值，我们对这个均值分两步处理，首先给出一个初始的均值，再根据状态方程对其下一时刻的位置进行预测，之后再通过观测方程对均值进行更新，为了描述实际的位置和估计的位置之间的偏差，引入协方差矩阵，同时对协方差矩阵进行预测和更新，不断迭代预测和更新的过程，即不断对随机变量的均值和协方差进行参数估计，即可由估计的均值和协方差得到最终的卡尔曼滤波结果。

卡尔曼滤波模型的描述

动态模型为了描述卡尔曼滤波模型，我们引入动态模型：



其中，x_{t-1}, x_t, x_{t+1} 为系统分别在 $t-1, t, t+1$ 时刻的实际状态，y_{t-1}, y_t, y_{t+1} 为系统在 $t-1, t, t+1$ 时刻的观测值。

定义：$P\{x_t|x_{t-1}\}$为系统状态的转移概率。

$P\{y_t|x_t\}$为系统状态的发射概率。

卡尔曼滤波模型的描述

线性高斯动态模型因为卡尔曼滤波是对于**线性高斯动态模型**的滤波，因此需要从线性高斯过程来研究卡尔曼滤波模型。对于一个**线性高斯动态模型**，其:

$$\begin{aligned} \text{过程函数：} & x_t = Ax_{t-1} + B + w_t \quad & w_t \sim \mathcal{N}(0, Q_t) \\ \text{观测函数：} & y_t = Hx_t + C + v_t \quad & v_t \sim \mathcal{N}(0, R_t) \end{aligned}$$

由于发射概率和转移概率都服从高斯分布，所以以上式就等价于

$$\begin{aligned} \text{转移概率：} & P\{x_t|x_{t-1}\} \sim \mathcal{N}(Ax_{t-1} + B, Q_t) \\ \text{发射概率：} & P\{y_t|x_t\} \sim \mathcal{N}(Hx_t, R_t) \end{aligned}$$

故求解动态模型的问题，就转化为求解参数向量 $[A\ B\ Q_t\ H\ C\ R_t]$ 的问题。

卡尔曼滤波模型的描述

由于卡尔曼滤波是滤波，因此根据滤波的定义，求解条件概率 $P(x_t|y_1......y_t)$ 即可实现滤波。根据乘法公式：

$$\begin{aligned} P(x_t|y_1......y_t) &\propto P(x_t y_1......y_t) \\ &= P(y_t|x_t y_1......y_{t_1})P(x_t y_1......y_{t-1}) \end{aligned}$$

因为 y_t 与 $y_1, y_2, ..., y_{t-1}$ 之间相互独立，上式可进一步化为：

$$\begin{aligned} P(x_t|y_1 y_2...y_t) &= P(y_t|x_t)P(x_t y_1 y_2...y_{t-1}) \\ &= P(y_1|x_t)P(x_t|y_1 y_2...y_{t-1})P(y_1 y_2...y_{t-1}) \end{aligned}$$

因为观测向量的概率 $P(y_1 Y_2...y_{t-1})$ 为常量，可得

$$P(x_t|y_1 y_2...y_t) \propto P(y_t|x_t)P(x_t|y_1 y_2...y_{t-1})$$

其中，$P(y_t|x_t)$ 即为系统的发射概率，而 $P(x_t|y_1, y_2, ..., y_{t-1})$ 则可以描述为由 $\vec{y} = [y_1\ y_2\ ...\ y_{t-1}]$ 对 x_t 的预测模型，即已知 $t-1$ 时刻与之前时刻的观测值，预测 t 时刻的实际状态。

卡尔曼滤波的求解

由 $P(x_t|y_1......y_t) \propto P(y_t|x_t)P(x_t|y_1......y_{t-1})$ 可知，求解卡尔曼滤波问题即求解 $P(x_t|y_1......y_{t-1})$

对 t 时刻的状态进行分解，即对 x_{t-1} 的状态进行遍历：

$$\begin{aligned}
P(x_t|y_1......y_{t-1}) &= \int_{x_{t-1}} P(x_t, x_{t-1}|y_1......y_{t-1})dx_{t-1} \\
&= \int_{x_{t-1}} P(x_t|x_{t-1}y_1......y_{t-1})P(x_{t-1}|y_1......y_{t-1})dx_{t-1} \\
&\quad \text{由于 } x_t \text{ 与 } y_1...y_t \text{ 无关} \\
&= \int_{x_{t-1}} P(x_t|x_{t-1})P(x_{t-1}|y_1......y_{t-1})dx_{t-1}
\end{aligned}$$

由乘法公式：

$$P(x_t|y_1......y_t) \propto P(y_t|x_t) \int_{x_{t-1}} P(x_t|x_{t-1})P(x_{t-1}|y_1......y_{t-1})dx_{t-1}$$

由上式形式可以看出 $P(x_t|y_1......y_t)$ 为 t 时刻的滤波结果，$P(x_{t-1}|y_1......y_{t-1})$ 是 $t-1$ 时刻的滤波结果，所以卡尔曼滤波是一个不断递归迭代的过程，每一次的滤波结果可根据转移概率和上一次滤波结果得出。

卡尔曼滤波的求解

由于卡尔曼滤波是不断递归迭代的过程，因此其求解过程可分为两个步骤：

1. 预测：根据 $y_1, y_2, ..., y_{t-1}$ 预测 x_t 的状态, 即求解 $P\{x_t|y_1 y_2 ... y_{t-1}\}$。
2. 更新：根据 $y_1, y_2, ..., y_{t-1}$ 和 y_t 对 x_t 的状态进行更新，即对 x_t 的滤波。

根据高斯线性系统的性质，可得：

$$\begin{aligned} \text{预测的概率：} P(x_t \mid y_1 \dots y_{t-1}) &= \mathcal{N}(\bar{\mu}_t, \bar{\Sigma}_t) \\ \text{更新的概率：} P(x_t \mid y_1 \dots y_t) &= \mathcal{N}(\hat{\mu}_t, \hat{\Sigma}_t) \end{aligned}$$

即卡尔曼滤波问题的求解转化为不断的代入预测更新方程求解 $\hat{\mu}_t, \hat{\Sigma}_t, \bar{\mu}_t, \bar{\Sigma}_t$ 这四个参数的问题。其递归流程如下：

$$\begin{array}{llll} t=1, & P(x_1|y_1) & \sim \mathcal{N}(\hat{\mu}_1, \hat{\Sigma}_1) & \text{(更新)} \\ t=2, & P(x_2|y_1) & \sim \mathcal{N}(\bar{\mu}_2, \bar{\Sigma}_2) & \text{(预测)} \\ & P(x_2|y_1, y_2) & \sim \mathcal{N}(\hat{\mu}_2, \hat{\Sigma}_2) & \text{(更新)} \\ t=3, & P(x_3|y_1, y_2) & \sim \mathcal{N}(\bar{\mu}_3, \bar{\Sigma}_3) & \text{(预测)} \\ & P(x_3|y_1, y_2, y_3) & \sim \mathcal{N}(\hat{\mu}_3, \hat{\Sigma}_3) & \text{(更新)} \end{array}$$

$$\begin{array}{llll} \vdots & & \vdots & \vdots \\ t=t, & P(x_t|y_1, ..., y_{t-1}) & \sim \mathcal{N}(\bar{\mu}_t, \bar{\Sigma}_t) & \text{(预测)} \\ & P(x_t|y_1, ..., y_t) & \sim \mathcal{N}(\hat{\mu}_t, \hat{\Sigma}_t) & \text{(更新)} \end{array}$$

不断代入公式求解 $\hat{\mu}_t, \hat{\Sigma}_t, \bar{\mu}_t, \bar{\Sigma}_t$

卡尔曼滤波的求解

因为 $P(x_{t-1} \mid y_1 \ldots y_{t-1}) \sim N(\hat{\mu}_{t-1}, \hat{\Sigma}_{t-1})$，因此可将其改写为常数叠加高斯变量的条件随机变量形式 $x_{t-1} \mid y_1 \ldots y_{t-1} = E[x_{t-1}] + \triangle x_{t-1}$，即在给定观测值 $y_1, y_2, ..., y_{t-1}$ 条件下的随机变量 x_t 的值，其中 $\triangle x_{t-1} \sim N(0, \hat{\Sigma}_{t-1})$
代入线性高斯动态模型即得

$$\begin{aligned} x_t \mid y_1 \ldots y_{t-1} &= Ax_{t-1} + w = A(E[x_{t-1}] + \triangle x_{t-1}) + w \\ &= AE[x_{t-1}] + A\triangle x_{t-1} + w \\ &\circeq E[x_t] + \triangle x_t \end{aligned}$$

即 x_t 在给定 $y_1 ... y_t$ 的条件下也是一个高斯随机变量，其均值为 $AE[x_{t-1}]$，协方差矩阵为 $E[\Delta x_t \Delta {x_t}^T]$
同理可得

$$\begin{aligned} P(y_t \mid y_1 \ldots y_{t-1}) &= Hx_t + v = H(AE[x_{t-1} + A\triangle x_{t-1} + w]) + v \\ &= HAE[x_{t-1}] + HA\triangle x_{t-1} + Hw + v \\ &= E[y_t] + \triangle y_t \end{aligned}$$

$$\begin{aligned} \text{故有：}\ & P(x_t \mid y_1 \ldots y_{t-1}) \sim N(AE[x_{t-1}], \triangle x_t \triangle x_t^T) \\ & P(y_t \mid y_1 \ldots y_{t-1}) \sim N(HAE[x_{t-1}], \triangle y_t \triangle y_t^T) \end{aligned}$$

卡尔曼滤波的求解

为了求解条件概率 $P(x_t|y_1...y_t)$ 和 $P(y_t|y_1...y_t)$，我们求其联合概率分布 $P(x_t,y_t|y_1...y_t)$
因为 x_t 与 y_t 都服从条件高斯分布，因此其联合概率：

$$P(x_t,y_t|y_1...y_{t-1}) = N\left(\begin{bmatrix} AE[x_{t-1}] \\ HAE[x_{t-1}] \end{bmatrix} \begin{bmatrix} E[\Delta x_t \Delta {x_t}^T] & E[\Delta x_t \Delta {y_t}^T] \\ E[\Delta y_t \Delta {x_t}^T]^T & E[\Delta y_t \Delta {y_t}^T] \end{bmatrix}\right)$$

因此，求解卡尔曼滤波的问题转化为对协方差矩阵

$$K = \begin{bmatrix} E[\Delta x_t \Delta {x_t}^T] & E[\Delta x_t \Delta {y_t}^T] \\ E[\Delta y_t \Delta {x_t}^T]^T & E[\Delta y_t \Delta {y_t}^T] \end{bmatrix}$$

的求解问题。

卡尔曼滤波的求解

依次化简矩阵元素可得：

$$
\begin{aligned}
E[\Delta x_t {\Delta x_t}^T] &= E[(A\Delta x_{t-1} + w)(A\Delta x_{t-1} + w)^T] \\
&= E[A\Delta x_{t-1}\Delta x_{t-1}^T A^T] + E[ww^T] \\
&= A\bar{\Sigma}_{t-1}A^T + Q_t
\end{aligned}
$$

$$
\begin{aligned}
E[\Delta y_t {\Delta x_t}^T] &= E[(HA\Delta x_{t-1} + Hw + v)(A\Delta x_{t-1} + w)^T] \\
&= E[HA\Delta x_{t-1}\Delta x_{t-1}A^T + Hww^T] \\
&= HA\bar{\Sigma}_{t-1}A_T + HQ_t
\end{aligned}
$$

$$
\begin{aligned}
E[\Delta x_t {\Delta y_t}^T] &= E[(A\Delta x_{t-1} + w)(HA\Delta x_{t-1} + Hw + v)^T] \\
&= E[HA\Delta x_{t-1}{\Delta x_{t-1}}^T A^T H^T + ww^T H^T] \\
&= A\bar{\Sigma}_{t-1}A^T H^T + Q_t H^T
\end{aligned}
$$

$$
\begin{aligned}
E[\Delta y \Delta y^T] &= = E[(HA\Delta x_{t-1} + Hw + v)(HA\Delta x_{t-1} + Hw + v)^T] \\
&= E[HA\Delta x_{t-1}{\Delta x_{t-1}}^T A^T H^T + Hww^T H^T + vv^T] \\
&= HA\bar{\Sigma}_{t-1}A^T H^T + HQ_t H^T + R_t
\end{aligned}
$$

卡尔曼滤波的求解

由上述化简可得:

$$
\begin{aligned}
&P(x_t, y_t|y_1, y_2, ..., y_{t-1})\\
=&\mathcal{N}\left(\begin{bmatrix} A\bar{\mu}_{t-1} \\ HA\bar{\mu}_{t-1}A^T \end{bmatrix} \begin{bmatrix} A\bar{\Sigma}_{t-1} + Q_t & [A\bar{\Sigma}_{t-1}A^T + Q_t]H^T \\ H[A\bar{\Sigma}_{t-1}A^T + Q_t] & H[A\bar{\Sigma}_{t-1}A^T + Q_t]H^T + R_t \end{bmatrix}\right)
\end{aligned}
$$

通常，对于服从二维高斯分布的随机变量 u 和 v，有:

$$
\begin{bmatrix} u \\ v \end{bmatrix} = \mathcal{N}\left(\begin{bmatrix} \mu_u \\ \mu_v \end{bmatrix} \begin{bmatrix} \Sigma_u & {\Sigma_{uv}}^T \\ \Sigma_{uv} & {\Sigma_v}^T \end{bmatrix}\right)
$$

我们可以求出其条件概率:

$$
P(u|v) = \mathcal{N}\left(\mu_u + \Sigma_{uv}{\Sigma_v}^{-1}(v - \mu_v), \Sigma_u - \Sigma_{uv}\Sigma_v^{-1}{\Sigma_{uv}}^T\right)
$$

即由联合概率 $P(x_t, y_t|y_1, y_2, ..., y_{t-1})$ 可以求得条件概率:

$$
P(x_t|y_1, y_2, ..., y_t) = \mathcal{N}\left(\mu_{x_t} + \Sigma_{x_t y_t}{\Sigma_{y_t}}^{-1}(y_t - \mu_{y_t}), \Sigma_{x_t} - \Sigma_{x_t y_t}\Sigma_{y_t}^{-1}{\Sigma_{x_t y_t}}^T\right)
$$

由上式即可得到卡尔曼滤波的滤波结果。

卡尔曼滤波的求解

根据所建立的状态方程和更新的概率分布 $P(x_t|y_1,,...,y_t)$，在给定系统初始状态的情况看下，可得如下的迭代过程：

$$\begin{pmatrix} \hat{\mu}_1 \\ \hat{\Sigma}_1 \end{pmatrix} \xrightarrow{x_2=Ax_1+B+w_2} \begin{pmatrix} \bar{\mu}_2 \\ \bar{\Sigma}_2 \end{pmatrix} \xrightarrow{P(x_2|y_1,y_2)\propto P(y_2|x_2)P(x_2|y_1)} \begin{pmatrix} \hat{\mu}_2 \\ \hat{\Sigma}_2 \end{pmatrix}$$

$$\cdots$$

$$\begin{pmatrix} \hat{\mu}_{t-1} \\ \hat{\Sigma}_{t-1} \end{pmatrix} \xrightarrow{x_t=Ax_{t-1}+B+w_t} \begin{pmatrix} \bar{\mu}_t \\ \bar{\Sigma}_t \end{pmatrix} \xrightarrow{P(x_t|y_1,\ldots,y_t)\propto P(y_t|x_t)P(x_t|y_1,\ldots,y_{t-1})} \begin{pmatrix} \hat{\mu}_t \\ \hat{\Sigma}_t \end{pmatrix}$$

由以上迭代过程即可不断的更新均值和协方差参数 $\hat{\mu}_t, \hat{\Sigma}_t, \bar{\mu}_t, \bar{\Sigma}_t$，并不断的由这些参数预测下一个时刻的系统状态 x_t。