0. 学习动机

方舱医院中为了改善cartographer定位会出现的跳变现象，使用IMU进行航迹推算，将它们的定位结果进行简单的融合，当cartographer定位和IMU推算的结果差距不大时，IMU使用cartographer的结果，重新开始航迹推算，如果差距过大，选择相信更加稳定的IMU结果，用IMU的结果在地图中进行重定位；

可以改善地图定位的跳变现象，但是当地图定位慢慢发生偏移的时候，由于设置阈值的关系，IMU会被地图定位带偏，所以，为了更好的对定位结果进行融合，需要使用滤波器的方法。

与此同时，为了实现新型的地面分割算法，也需要滤波器的知识，所以学习以卡尔曼滤波为代表的滤波方法。

学习方式为：

1.【网课】深蓝学院-多传感器融合定位-第四章；

2.【博客】

0.1 通俗理解卡尔曼滤波

0.1.1

假设你有两个传感器，测的是同一个信号。可是它们每次的读数都不太一样，怎么办？

取平均。

再假设你知道其中贵的那个传感器应该准一些，便宜的那个应该差一些。那有比取平均更好的办法吗？

加权平均。

怎么加权？假设两个传感器的误差都符合正态分布，假设你知道这两个正态分布的方差，用这两个方差值，（此处省略若干数学公式），你可以得到一个“最优”的权重。

接下来，重点来了：假设你只有一个传感器，但是你还有一个数学模型。模型可以帮你算出一个值，但也不是那么准。怎么办？

把模型算出来的值，和传感器测出的值，（就像两个传感器那样），取加权平均。

OK，最后一点说明：你的模型其实只是一个步长的，也就是说，知道x(k)，我可以求x(k+1)。问题是x(k)是多少呢？答案：x(k)就是你上一步卡尔曼滤波得到的、所谓加权平均之后的那个、对x在k时刻的最佳估计值。

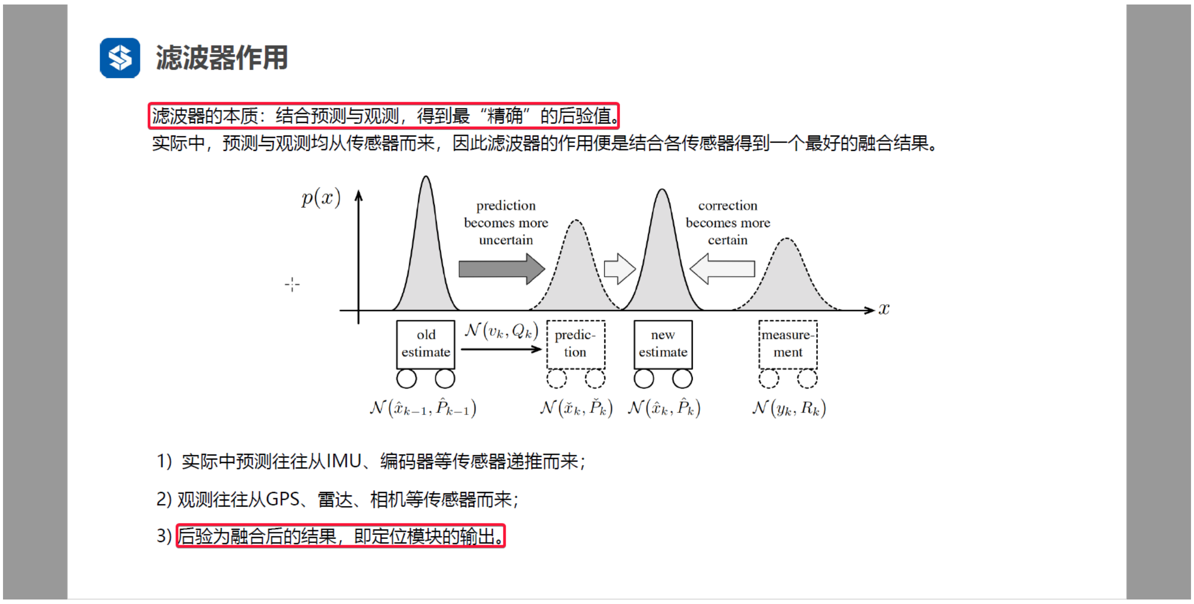
于是**迭代**也有了。

这就是卡尔曼滤波。

（无公式）

作者：Kent Zeng
链接：https://www.zhihu.com/question/23971601/answer/26254459
来源：知乎

1. 滤波器作用



融合直观的解释：

点云匹配：观测，波动比较大

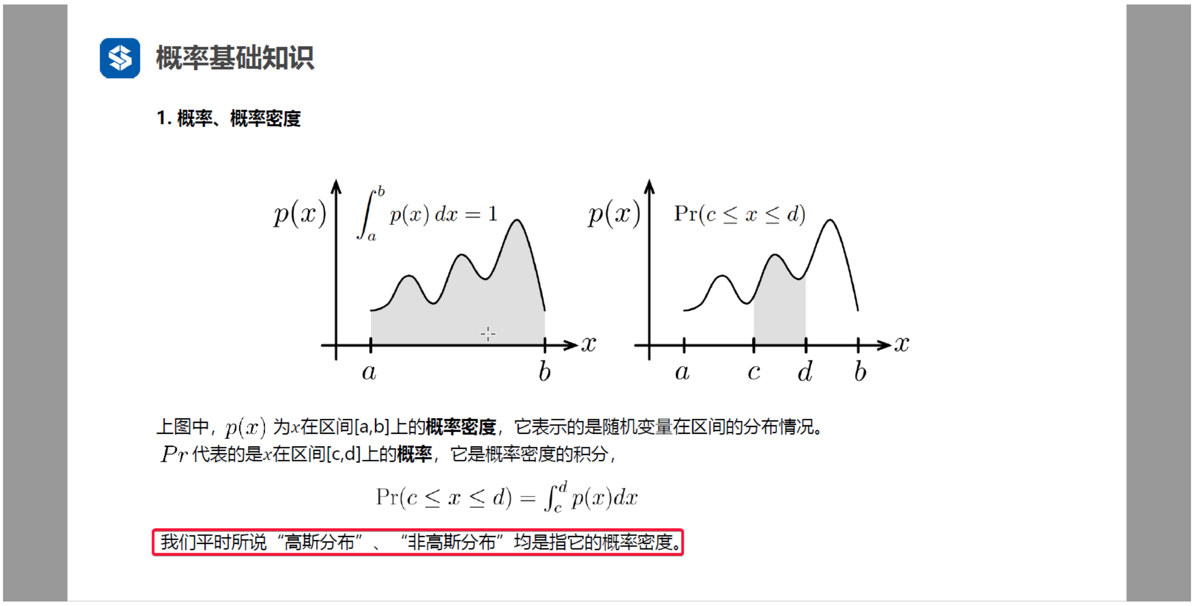
IMU：比较稳定，可以作为预测，但是有累积误差

可以使用点云匹配的结果对IMU进行修正

最后得到一个没有累积误差而且波动比较小的结果。

2. 概率基础知识

2.1 概率、概率密度



概率密度直观上就是概率的形状，表示在区间的分布情况，上图中，一个值落在[a,b]的概率是100%，但是落在中间某两个值之间的概率就需要使用概率密度函数来计算；

2.2 联合概率与条件概率

2. 联合概率

$x \in [a,b]$和 $y \in [r,s]$的联合概率密度函数可以表示为$p(x,y)$，其积分表示x和y同时处在某个区间的概率，满足下式：

$$\int_a^b \underbrace{\boxed{\int_r^s p(x,y)dy}}_{=p(x)} dx = 1$$

特别地，当x和y统计独立的时候，有

$$p(x,y) = p(x)p(y)$$

3. 条件概率

x关于y的条件概率密度函数可以表示为

$$p(x \mid y)$$

其含义是，在 $y \in [r,s]$ 的前提下，$x \in [a,b]$ 的概率分布，并且满足下式

$$\boxed{p(x) = \int_r^s p(x \mid y)p(y)dy}$$

特别地，当x和y统计独立的时候，有

$$p(x \mid y) = p(x)$$

条件概率p(x|y)：在y发生下x发生的概率

如果x，y统计独立，即x，y没有关系，y对x没有影响，那么，p(x|y)=p(x)

如果y对x有影响，y取任何一个值，x都有一个概率p(x)，为了得到完整的p(x),对y进行积分。

2.3 贝叶斯公式/推断

联合概率分解成条件概率和边缘概率的乘积，即

$$p(x,y) = p(x \mid y)p(y) = p(y \mid x)p(x)$$

重新整理，即可得贝叶斯公式

$$p(x \mid y) = \frac{p(y|x)p(x)}{p(y)}$$

贝叶斯推断可以理解为贝叶斯公式的运用，它是指，如果已知**先验**概率密度函数 $p(x)$，以及**传感器模型** $p(y \mid x)$，那么就可以根据贝叶斯公式**推断**出后验概率密度。

$$p(x \mid y) = \frac{p(y|x)p(x)}{\int p(y|x)p(x)\mathrm{d}x}$$

实际中，贝叶斯推断有时也叫贝叶斯估计。

贝叶斯估计具有**非常重大的意义**：

已知状态先验->可以推断出后验概率，也就是状态估计

比如说：

以下雨为一个状态x，当下雨时有乌云y的概率为先验p(y|x),使用贝叶斯公式可以得到p(x|y)，即，当有乌云时，会下雨的概率，这就是对状态x的估计。

2.4 高斯概率密度函数

2.4.1 高斯概率密度函数

概率基础知识

6. 高斯概率密度函数

一维情况下，高斯概率密度函数表示为：

$$p\left(x \mid \mu, \sigma^{2}\right)=\frac{1}{\sqrt{2 \pi \sigma^{2}}} \exp \left(-\frac{1}{2} \frac{(x-\mu)^{2}}{\sigma^{2}}\right)$$

其中 μ 为均值，σ^2 为方差。

多维情况下，高斯概率密度函数表示为

$$p(\boldsymbol{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})=\frac{1}{\sqrt{(2 \pi)^{N} \operatorname{det} \boldsymbol{\Sigma}}} \exp \left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}} \boxed{\boldsymbol{\Sigma}^{-1}}(\boldsymbol{x}-\boldsymbol{\mu})\right)$$

向量 矩阵

其中均值为 $\boldsymbol{\mu}$，方差为 $\boldsymbol{\Sigma}$。

一般把高斯分布写成 $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$。

卡尔曼滤波以高斯概率函数为基础

2.4.2 联合高斯密度函数

若有高斯分布

$$p(\boldsymbol{x})=\mathcal{N}\left(\boldsymbol{\mu}_{x}, \boldsymbol{\Sigma}_{xx}\right)$$
$$p(\boldsymbol{y})=\mathcal{N}\left(\boldsymbol{\mu}_{y}, \boldsymbol{\Sigma}_{yy}\right)$$

则它们的联合概率密度函数可以表示为

$$p(\boldsymbol{x}, \boldsymbol{y})=\mathcal{N}\left(\left[\begin{array}{c}\boldsymbol{\mu}_{x} \\ \boldsymbol{\mu}_{y}\end{array}\right],\left[\begin{array}{cc}\boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{xy} \\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{yy}\end{array}\right]\right)$$

由于联合概率满足下式

$$p(\boldsymbol{x}, \boldsymbol{y})=p(\boldsymbol{x} \mid \boldsymbol{y}) p(\boldsymbol{y})$$

该式在高斯分布的前提下可以重新分解。

由于高斯分布中指数项包含方差的求逆，而此处联合概率的方差是一个高维矩阵，对它求逆的简洁办法是运用舒尔补。

舒尔补的主要目的是把矩阵分解成上三角矩阵、对角阵、下三角矩阵乘积的形式，方便运算，即

$$\begin{aligned}&\left[\begin{array}{ll}\mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D}\end{array}\right] \\ =&\left[\begin{array}{cc}\mathbf{I} & \mathbf{B D}^{-1} \\ \mathbf{0} & \mathbf{I}\end{array}\right]\left[\begin{array}{cc}\Delta \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}\end{array}\right]\left[\begin{array}{cc}\mathbf{I} & \mathbf{0} \\ \mathbf{D}^{-1} \mathbf{C} & \mathbf{I}\end{array}\right]\end{aligned}$$

其中 $\Delta \mathbf{D}=\mathbf{A}-\mathbf{B D}^{-1} \mathbf{C}$ 称为矩阵D关于原矩阵的舒尔补。

此时有

$$\begin{aligned}&\left[\begin{array}{ll}\mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D}\end{array}\right]^{-1} \\ =&\left[\begin{array}{cc}\mathbf{I} & \mathbf{0} \\ -\mathbf{D}^{-1} \mathbf{C} & \mathbf{I}\end{array}\right]\left[\begin{array}{cc}\Delta_{\mathbf{D}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^{-1}\end{array}\right]\left[\begin{array}{cc}\mathbf{I} & -\mathbf{B D}^{-1} \\ \mathbf{0} & \mathbf{I}\end{array}\right]\end{aligned}$$

方便求逆

利用舒尔补，联合分布的方差矩阵可以写为

$$\left[\begin{array}{cc}\boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{xy} \\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{yy}\end{array}\right]=\left[\begin{array}{cc}\mathbf{1} & \boldsymbol{\Sigma}_{xy} \boldsymbol{\Sigma}_{yy}^{-1} \\ \mathbf{0} & \mathbf{1}\end{array}\right]\left[\begin{array}{cc}\boldsymbol{\Sigma}_{xx}-\boldsymbol{\Sigma}_{xy} \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{\Sigma}_{yx} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{yy}\end{array}\right]\left[\begin{array}{cc}\mathbf{1} & \mathbf{0} \\ \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{\Sigma}_{yx} & \mathbf{1}\end{array}\right]$$

它的逆矩阵为

$$\left[\begin{array}{cc}\boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{xy} \\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{yy}\end{array}\right]^{-1}=\left[\begin{array}{cc}\mathbf{1} & \mathbf{0} \\ -\boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{\Sigma}_{yx} & \mathbf{1}\end{array}\right]\left[\begin{array}{cc}\left(\boldsymbol{\Sigma}_{xx}-\boldsymbol{\Sigma}_{xy} \boldsymbol{\Sigma}_{yy}^{-1} \boldsymbol{\Sigma}_{yx}\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{yy}^{-1}\end{array}\right]\left[\begin{array}{cc}\mathbf{1} & -\boldsymbol{\Sigma}_{xy} \boldsymbol{\Sigma}_{yy}^{-1} \\ \mathbf{0} & \mathbf{1}\end{array}\right]$$

联合分布 $p(\boldsymbol{x}, \boldsymbol{y})$ 仍为高斯分布，

$$p(\boldsymbol{x}, \boldsymbol{y})=\mathcal{N}\left(\left[\begin{array}{l}\boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y\end{array}\right],\left[\begin{array}{ll}\boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{xy} \\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{yy}\end{array}\right]\right)$$

它的指数部分的二次项包含如下内容

$$\begin{aligned}
&\left(\left[\begin{array}{l}\boldsymbol{x} \\ \boldsymbol{y}\end{array}\right]-\left[\begin{array}{l}\boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y\end{array}\right]\right)^{\mathrm{T}}\left[\begin{array}{ll}\boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{xy} \\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{yy}\end{array}\right]^{-1}\left(\left[\begin{array}{l}\boldsymbol{x} \\ \boldsymbol{y}\end{array}\right]-\left[\begin{array}{l}\boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y\end{array}\right]\right) \\
=&\left(\left[\begin{array}{l}\boldsymbol{x} \\ \boldsymbol{y}\end{array}\right]-\left[\begin{array}{l}\boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y\end{array}\right]\right)^{\mathrm{T}}\left[\begin{array}{cc}\mathbf{1} & \mathbf{0} \\ -\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx} & \mathbf{1}\end{array}\right]\left[\begin{array}{cc}\left(\boldsymbol{\Sigma}_{xx}-\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx}\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{yy}^{-1}\end{array}\right]\left[\begin{array}{cc}\mathbf{1} & -\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1} \\ \mathbf{0} & \mathbf{1}\end{array}\right]\left(\left[\begin{array}{l}\boldsymbol{x} \\ \boldsymbol{y}\end{array}\right]-\left[\begin{array}{l}\boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y\end{array}\right]\right) \\
=&\boxed{\left(\boldsymbol{x}-\boldsymbol{\mu}_x-\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\left(\boldsymbol{y}-\boldsymbol{\mu}_y\right)\right)^{\mathrm{T}}\left(\boldsymbol{\Sigma}_{xx}-\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx}\right)^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}_x-\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\left(\boldsymbol{y}-\boldsymbol{\mu}_y\right)\right)} \\
&+\boxed{\left(\boldsymbol{y}-\boldsymbol{\mu}_y\right)^{\mathrm{T}}\boldsymbol{\Sigma}_{yy}^{-1}\left(\boldsymbol{y}-\boldsymbol{\mu}_y\right)}
\end{aligned}$$

最后得到两个二次项的和，由于同底数幂相乘后，底数不变，指数相加，且 $\boxed{p(\boldsymbol{y})=\mathcal{N}\left(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_{yy}\right)}$

因此有 $\boxed{p(\boldsymbol{x} \mid \boldsymbol{y})=\mathcal{N}\left(\boldsymbol{\mu}_x+\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\left(\boldsymbol{y}-\boldsymbol{\mu}_y\right), \boldsymbol{\Sigma}_{xx}-\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx}\right)}$

(包括以后）最终目的就是为了获得p(x|y)

2.4.3 高斯随机变量的线性分布

在上面的例子中，若已知x和y之间有如下关系

$$\boldsymbol{y}=\boldsymbol{G}\boldsymbol{x}+\boldsymbol{n}$$

其中G是一个常量矩阵，$\boldsymbol{n}=\mathcal{N}(\mathbf{0}, \boldsymbol{R})$为零均值白噪声，在实际中指的是观测噪声。则$x$和$y$的均值和方差之间必然存在联系，其联系可通过以下推导获得。

均值

$$\begin{aligned}
\boldsymbol{\mu}_y &= E[\boldsymbol{y}] \\
&= E[\boldsymbol{G}\boldsymbol{x}+\boldsymbol{n}] \\
&= \boldsymbol{G}E[\boldsymbol{x}]+E[\boldsymbol{n}] \\
&= \boldsymbol{G}\boldsymbol{\mu}_x
\end{aligned}$$

方差

$$\begin{aligned}
\boldsymbol{\Sigma}_{yy} &= \boldsymbol{\Sigma}(\boldsymbol{G}\boldsymbol{x})+\boldsymbol{\Sigma}(\boldsymbol{n}) \\
&= E\left[\left(\boldsymbol{G}\boldsymbol{x}-\boldsymbol{\mu}_y\right)\left(\boldsymbol{G}\boldsymbol{x}-\boldsymbol{\mu}_y\right)^{\mathrm{T}}\right]+\boldsymbol{R} \\
&= \boldsymbol{G}E\left[\left(\boldsymbol{x}-\boldsymbol{\mu}_x\right)\left(\boldsymbol{x}-\boldsymbol{\mu}_x\right)^{\mathrm{T}}\right]\boldsymbol{G}^{\mathrm{T}}+\boldsymbol{R} \\
&= \boldsymbol{G}\boldsymbol{\Sigma}_{xx}\boldsymbol{G}^{\mathrm{T}}+\boldsymbol{R}
\end{aligned}$$

方差的交叉项

$$\begin{aligned}
\boldsymbol{\Sigma}_{xy} &= E\left[\left(\boldsymbol{x}-\boldsymbol{\mu}_x\right)\left(\boldsymbol{y}-\boldsymbol{\mu}_y\right)^{\mathrm{T}}\right] \\
&= E\left[\left(\boldsymbol{x}-\boldsymbol{\mu}_x\right)\left(\boldsymbol{G}\boldsymbol{x}-\boldsymbol{G}\boldsymbol{\mu}_x+\boldsymbol{n}\right)^{\mathrm{T}}\right] \\
&= E\left[\left(\boldsymbol{x}-\boldsymbol{\mu}_x\right)\left(\boldsymbol{G}\boldsymbol{x}-\boldsymbol{G}\boldsymbol{\mu}_x\right)^{\mathrm{T}}+\left(\boldsymbol{x}-\boldsymbol{\mu}_x\right)\boldsymbol{n}^{\mathrm{T}}\right] \\
&= \boldsymbol{\Sigma}_{xx}\boldsymbol{G}^{\mathrm{T}}+E\left[\left(\boldsymbol{x}-\boldsymbol{\mu}_x\right)\boldsymbol{n}^{\mathrm{T}}\right] \\
&= \boldsymbol{\Sigma}_{xx}\boldsymbol{G}^{\mathrm{T}}
\end{aligned}$$

同理可得 $\boldsymbol{\Sigma}_{yx}=\boldsymbol{\Sigma}_{xy}^{\mathrm{T}}=\boldsymbol{G}\boldsymbol{\Sigma}_{xx}$

3. 滤波器基本原理

3.1 状态估计模型

1. 状态估计模型

实际状态估计任务中，待估计的后验概率密度可以表示为：

$$p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k}\right)$$

其中

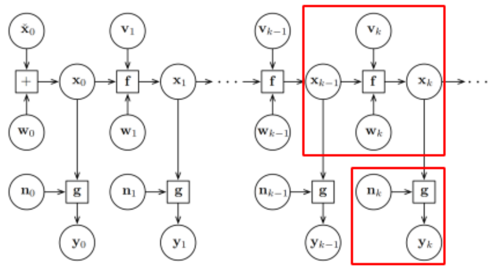
$\check{\boldsymbol{x}}_0$ 表示的是状态初始值

$\boldsymbol{v}_{1:k}$ 表示从1到k时刻的输入

$\boldsymbol{y}_{0:k}$ 表示从0到k时刻的观测

因此，滤波问题可以直观表示为，根据所有历史数据(输入、观测、初始状态)，得出最终的融合结果。

历史数据之间的关系，可以用下面的图模型表示，



图模型中体现了马尔可夫性，即当前状态只跟前一时刻状态相关，和其他历史时刻状态无关。数学表达该性质，

$$\boxed{\begin{array}{ll}\text{运动方程：} & \boldsymbol{x}_k=\boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{v}_k, \boldsymbol{w}_k\right) \\ \text{观测方程：} & \boldsymbol{y}_k=\boldsymbol{g}\left(\boldsymbol{x}_k, \boldsymbol{n}_k\right)\end{array}}$$

3.2 贝叶斯滤波

2. 贝叶斯滤波

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

根据贝叶斯公式，k时刻后验概率密度可以表示为

$$\begin{aligned} p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k}\right) &= \frac{p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k, \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right)}{p\left(\boldsymbol{y}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right)} \\ &= \eta p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k, \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) \end{aligned}$$

与x无关

$\downarrow$

η

根据观测方程，$\boldsymbol{y}_k$ 只和 $\boldsymbol{x}_k$ 相关，因此上式可以简写为

$$p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k}\right) = \eta p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k\right) p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right)$$

应用系统的马尔可夫性进一步化简公式，

$$\begin{aligned} & p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) \\ & = \int p\left(\boldsymbol{x}_k, \boldsymbol{x}_{k-1} \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) \mathrm{d}\boldsymbol{x}_{k-1} \\ & = \int p\left(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}, \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) p\left(\boldsymbol{x}_{k-1} \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) \mathrm{d}\boldsymbol{x}_{k-1} \\ & = \int p\left(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}, \boldsymbol{v}_k\right) p\left(\boldsymbol{x}_{k-1} \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) \mathrm{d}\boldsymbol{x}_{k-1} \end{aligned}$$

马尔可夫性：当前状态x_k只和前一时刻x_{k-1}有关

$$p(x) = \int p(x|y)p(y)dy \tag{1}$$

利用(1) 可以让任意其他变量y与x发生关系，将y变成x_{k-1},x_k就和x_{k-1}发生了关联.

根据观测方程，$\boldsymbol{y}_k$ 只和 $\boldsymbol{x}_k$ 相关，因此上式可以简写为

$$p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k}\right) = \eta p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k\right) p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right)$$

应用系统的马尔可夫性进一步化简公式，

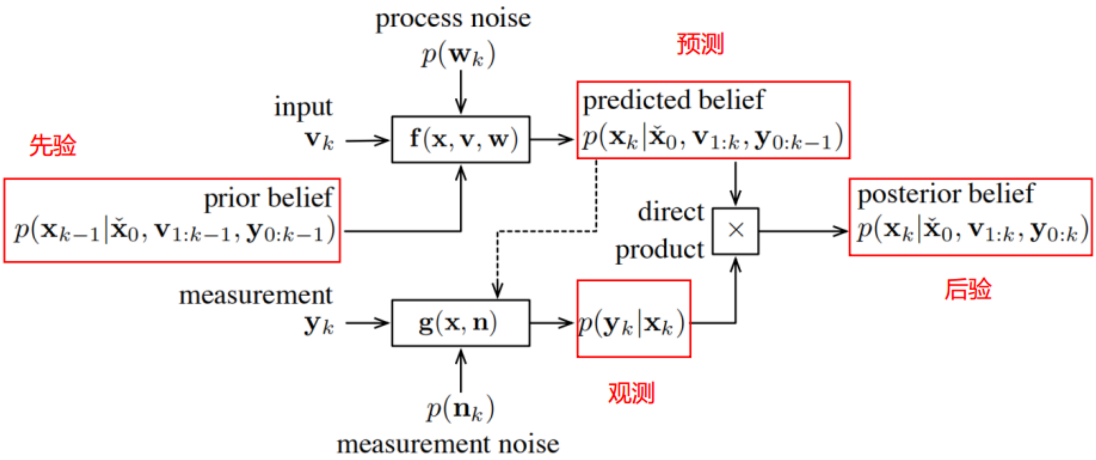
$$\begin{aligned} & p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) \\ & = \int p\left(\boldsymbol{x}_k, \boldsymbol{x}_{k-1} \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) \mathrm{d}\boldsymbol{x}_{k-1} \\ & = \int p\left(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}, \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) p\left(\boldsymbol{x}_{k-1} \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) \mathrm{d}\boldsymbol{x}_{k-1} \\ & = \int p\left(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}, \boldsymbol{v}_k\right) p\left(\boldsymbol{x}_{k-1} \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) \mathrm{d}\boldsymbol{x}_{k-1} \end{aligned}$$

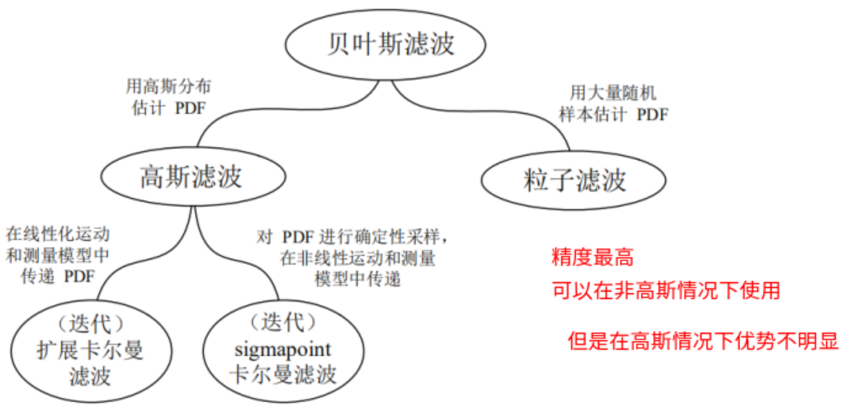
经过以上化简，最终后验概率可以写为

$$\underbrace{p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k}\right)}_{\text{estimate}} = \eta \boxed{\underbrace{p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k\right)}_{\boldsymbol{g}(\cdot)}} \int \boxed{\underbrace{p\left(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}, \boldsymbol{v}_k\right)}_{\boldsymbol{f}(\cdot)}} \boxed{\underbrace{p\left(\boldsymbol{x}_{k-1} \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k-1}, \boldsymbol{y}_{0:k-1}\right)}_{\text{last estimate}}} \mathrm{d}\boldsymbol{x}_{k-1}$$

观测　　　预测　　　先验

根据以上结果，可以画出贝叶斯滤波的信息流图如下



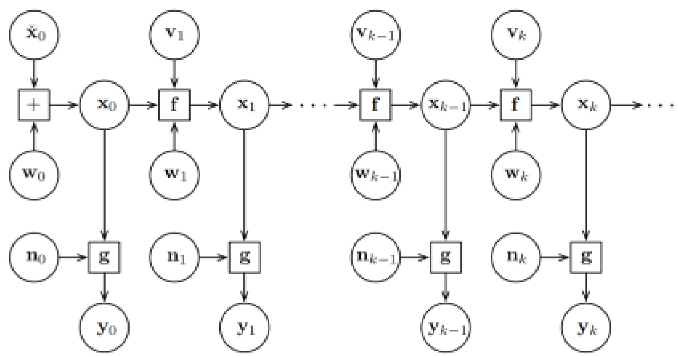


1）在高斯假设前提下，用贝叶斯滤波的原始形式比较复杂，可以利用高斯的特征得到简化形式，即广义高斯滤波，后面KF、EKF、IEKF、UKF的推导均采用这种形式，只有PF例外，因为它是针对非高斯的。

2) 实际中，粒子滤波多用于(早期的)2D激光SLAM方案中，不属于本门课程讲解重点，因此只做原理介绍。

3.3 卡尔曼滤波(KF)推导

历史数据之间的关系，可以用下面的图模型表示，



图模型中体现了马尔可夫性，即当前状态只跟前一时刻状态相关，和其他历史时刻状态无关。数学表达该性质，

$$\text{运动方程：}\quad \boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{v}_k, \boldsymbol{w}_k\right)$$

$$\text{观测方程：}\quad \boldsymbol{y}_k = \boldsymbol{g}\left(\boldsymbol{x}_k, \boldsymbol{n}_k\right)$$

在线性高斯假设下，上式可以重新写为下面的形式(为了和后面符号对应)

$$\begin{aligned} &\text{运动方程：} \quad \boldsymbol{x}_k = \boldsymbol{F}\left(\boldsymbol{x}_{k-1}, \boldsymbol{v}_k\right) + \boldsymbol{B}_{k-1} \boldsymbol{w}_k \\ &\text{观测方程：} \quad \boldsymbol{y}_k = \boldsymbol{G}\left(\boldsymbol{x}_k\right) + \boldsymbol{C}_k \boldsymbol{n}_k \end{aligned}$$

把上一时刻的后验状态写为

$$p\left(\boldsymbol{x}_{k-1} \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k-1}, \boldsymbol{y}_{0:k-1}\right) = \mathcal{N}\left(\hat{\boldsymbol{x}}_{k-1}, \hat{\boldsymbol{P}}_{k-1}\right)$$

则当前时刻的预测值为

$$\check{\boldsymbol{x}}_k = \boldsymbol{F}\left(\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{v}_k\right)$$

根据高斯分布的线性变化，它的方差为

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

其中$\boldsymbol{Q}_k$为当前输入噪声的方差。

线性假设下:

运动方程:

$$x_k = f(x_{k-1}, u_k) + \varepsilon_k = A_k x_{k-1} + B_k u_k + \varepsilon_k$$

观测方程:

$$z_k = h(x_k) + \delta_k = C_k x_k + \delta_k$$

若把k时刻状态和观测的联合高斯分布写为

$$p\left(\boldsymbol{x}_k, \boldsymbol{y}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k-1}\right) = \mathcal{N}\left(\begin{bmatrix} \boldsymbol{\mu}_{x,k} \\ \boldsymbol{\mu}_{y,k} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma}_{xx,k} & \boldsymbol{\Sigma}_{xy,k} \\ \boldsymbol{\Sigma}_{yx,k} & \boldsymbol{\Sigma}_{yy,k} \end{bmatrix}\right)$$

根据第1节7)中的推导结果，k时刻的后验概率可以写为

$$\begin{aligned} & p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k}\right) \\ & \quad = \mathcal{N}(\underbrace{\boldsymbol{\mu}_{x,k} + \boldsymbol{\Sigma}_{xy,k} \boldsymbol{\Sigma}_{yy,k}^{-1}\left(\boldsymbol{y}_k - \boldsymbol{\mu}_{y,k}\right)}_{\hat{\boldsymbol{x}}_k}, \underbrace{\boldsymbol{\Sigma}_{xx,k} - \boldsymbol{\Sigma}_{xy,k} \boldsymbol{\Sigma}_{yy,k}^{-1} \boldsymbol{\Sigma}_{yx,k}}_{\hat{P}_k}) \end{aligned}$$

其中 $\hat{\boldsymbol{x}}_k$和$\hat{\boldsymbol{P}}_k$分别为后验均值和方差。若定义

$$\boldsymbol{K}_k = \boldsymbol{\Sigma}_{xy,k} \boldsymbol{\Sigma}_{yy,k}^{-1}$$

则有

$$\begin{aligned} \hat{\boldsymbol{P}}_k &= \check{\boldsymbol{P}}_k - \boldsymbol{K}_k \boldsymbol{\Sigma}_{xy,k}^{\mathrm{T}} \\ \hat{\boldsymbol{x}}_k &= \check{\boldsymbol{x}}_k + \boldsymbol{K}_k\left(\boldsymbol{y}_k - \boldsymbol{\mu}_{y,k}\right) \end{aligned}$$

把第1节8)中的推导得出的线性变换后的均值、方差及交叉项带入上面的式子，可以得到：

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

$$\hat{\boldsymbol{P}}_k = \left(\mathbf{1} - \boldsymbol{K}_k \boldsymbol{G}_k\right) \check{\boldsymbol{P}}_k$$

$$\hat{\boldsymbol{x}}_k = \check{\boldsymbol{x}}_k + \boldsymbol{K}_k \left(\boldsymbol{y}_k - \boldsymbol{G}\left(\check{\boldsymbol{x}}_k\right)\right)$$

上面方程与之前所述预测方程(如下)，就构成了卡尔曼经典五个方程。

$$\check{\boldsymbol{x}}_k = \boldsymbol{F}\left(\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{v}_k\right)$$

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

需要说明的是，若不把第1节8)中的结果带入，而保留上页的原始形式，则对应的五个方程被称为广义高斯滤波。

………

1) 相比于推导方法，理解它的本质更重要，即“信息加权”和“权重调节”；

2) 滤波器还有其他很多种推导方式，只需理解它的一种推导方法即可，不要做“孔乙己”；

3) 实际工程中，滤波器的表现受多种因素的影响，现象千变万化，脱离工程不可能深入理解滤波器。

4) 在2D激光SLAM中，随着基于优化方案的优势越来越明显，基于PF的方法逐渐“过时”(尽管很多人并不赞同这一点)。

我认为不需要掌握卡尔曼滤波器的具体数学推导,只需要记住卡尔曼滤波的五个方程,理解物理意义,根据实际要求使用KF.

Algorithm Kalman_filter($\mu_{k-1}, \Sigma_{k-1}, u_k, z_k$):

for all x_k do

$$\bar{\mu}_k = A_k \mu_{k-1} + B_k u_k$$

$$\bar{\Sigma}_k = A_k \Sigma_{k-1} A_k^T + R_k$$

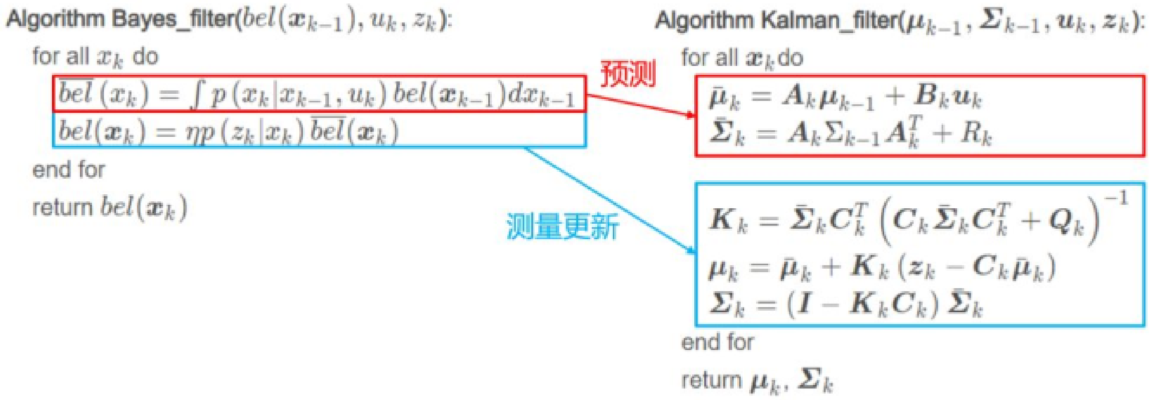
$$K_k = \bar{\Sigma}_k C_k^T \left(C_k \bar{\Sigma}_k C_k^T + Q_k\right)^{-1}$$

$$\mu_k = \bar{\mu}_k + K_k \left(z_k - C_k \bar{\mu}_k\right)$$

$$\Sigma_k = \left(I - K_k C_k\right) \bar{\Sigma}_k$$

end for

return μ_k, Σ_k



信息加权,权重调节:

卡尔曼滤波核心思想是将预测的状态和观测的状态进行加权融合,权重调节由卡尔曼增益K控制;

预测噪声方差R

观测噪声方差Q

谁越小,就越相信谁,体现在K: