深蓝学院
shenlanxueyuan.com

多传感器融合定位

第4章 基于滤波的融合方法

主讲人 任 乾

北京理工大学本硕
自动驾驶从业者

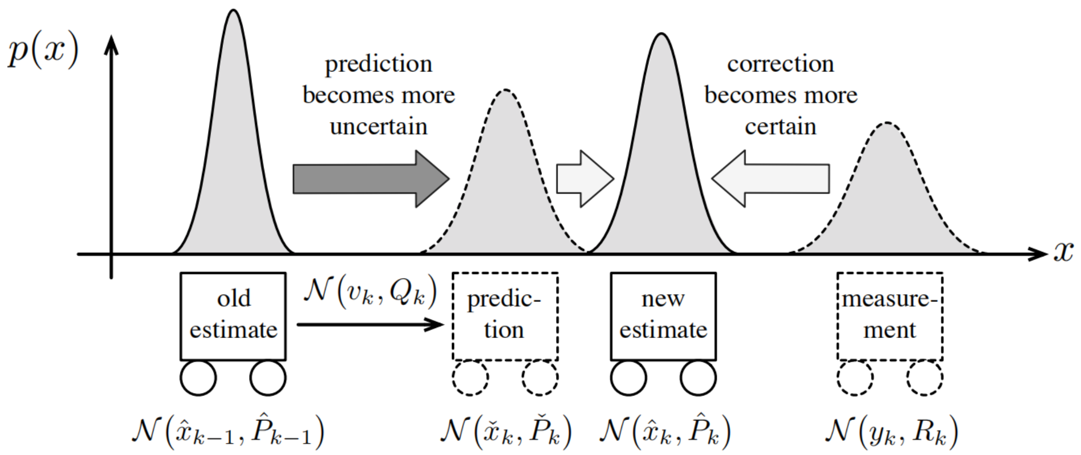
目录

目录

滤波器作用

滤波器的本质：结合预测与观测，得到最“精确”的后验值。

实际中，预测与观测均从传感器而来，因此滤波器的作用便是结合各传感器得到一个最好的融合结果。



1) 实际中预测往往从IMU、编码器等传感器递推而来；

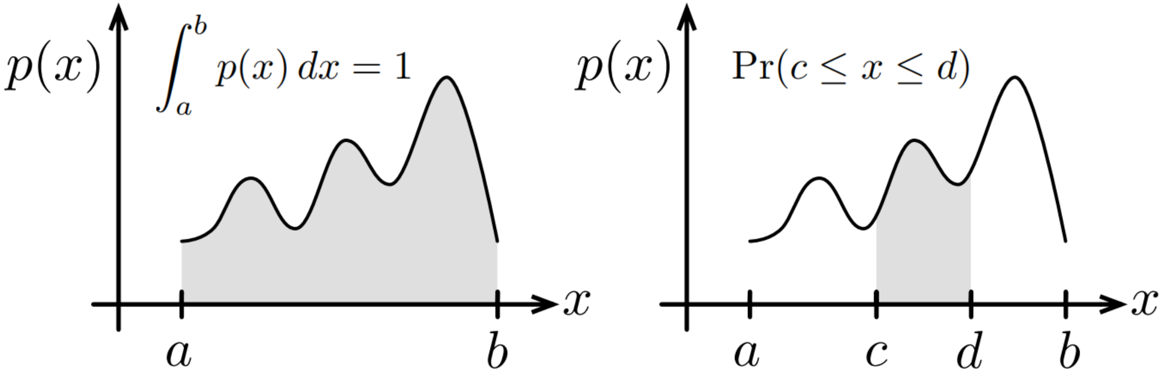
2) 观测往往从GPS、雷达、相机等传感器而来；

3) 后验为融合后的结果，即定位模块的输出。

目录

概率基础知识

1. 概率、概率密度



上图中，$p(x)$ 为x在区间[a,b]上的**概率密度**，它表示的是随机变量在区间的分布情况。

Pr 代表的是x在区间[c,d]上的**概率**，它是概率密度的积分，

$$\Pr(c \le x \le d) = \int_c^d p(x)dx$$

我们平时所说“高斯分布”、“非高斯分布”均是指它的概率密度。

概率基础知识

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：

$$\int_a^b \int_r^s p(x, y) dy dx = 1$$

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

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

特别地，当x和y统计独立的时候，有

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

概率基础知识

4. 贝叶斯公式

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

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

5. 贝叶斯推断

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。

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

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

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

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

7. 联合高斯概率密度函数

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

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

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

$$\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{bmatrix}$$
$$=\begin{bmatrix}\mathbf{I} & \mathbf{B}\mathbf{D}^{-1}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\Delta\mathbf{D} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}\end{bmatrix}\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{D}^{-1}\mathbf{C} & \mathbf{I}\end{bmatrix}$$

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

此时有

$$\begin{bmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{bmatrix}^{-1}$$
$$=\begin{bmatrix}\mathbf{I} & \mathbf{0}\\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}\end{bmatrix}\begin{bmatrix}\Delta_{\mathbf{D}}^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{D}^{-1}\end{bmatrix}\begin{bmatrix}\mathbf{I} & -\mathbf{B}\mathbf{D}^{-1}\\ \mathbf{0} & \mathbf{I}\end{bmatrix}$$

7. 联合高斯概率密度函数

利用舒尔补，联合分布的方差矩阵可以写为

$$\begin{bmatrix} \boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{xy} \\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{yy} \end{bmatrix} = \begin{bmatrix} \mathbf{1} & \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{xx} - \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx} & \mathbf{1} \end{bmatrix}$$

它的逆矩阵为

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

7. 联合高斯概率密度函数

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

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

它的指数部分的二次项包含如下内容

$$\begin{aligned}
&\left(\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{bmatrix}\right)^{\mathrm{T}} \begin{bmatrix} \boldsymbol{\Sigma}_{xx} & \boldsymbol{\Sigma}_{xy} \\ \boldsymbol{\Sigma}_{yx} & \boldsymbol{\Sigma}_{yy} \end{bmatrix}^{-1} \left(\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{bmatrix}\right) \\
&= \left(\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{bmatrix}\right)^{\mathrm{T}} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ -\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \left(\boldsymbol{\Sigma}_{xx} - \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx}\right)^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{yy}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{1} & -\boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \left(\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} - \begin{bmatrix} \boldsymbol{\mu}_x \\ \boldsymbol{\mu}_y \end{bmatrix}\right) \\
&= \left(\boldsymbol{x} - \boldsymbol{\mu}_x - \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\left(\boldsymbol{y} - \boldsymbol{\mu}_y\right)\right)^{\mathrm{T}} \left(\boldsymbol{\Sigma}_{xx} - \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx}\right)^{-1} \left(\boldsymbol{x} - \boldsymbol{\mu}_x - \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\left(\boldsymbol{y} - \boldsymbol{\mu}_y\right)\right) \\
&\quad + \left(\boldsymbol{y} - \boldsymbol{\mu}_y\right)^{\mathrm{T}} \boldsymbol{\Sigma}_{yy}^{-1} \left(\boldsymbol{y} - \boldsymbol{\mu}_y\right)
\end{aligned}$$

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

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

8. 高斯随机变量的线性分布

在上面的例子中，若已知$\boldsymbol{x}$和$\boldsymbol{y}$之间有如下关系

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

其中$\boldsymbol{G}$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$\boldsymbol{x}$和$\boldsymbol{y}$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 $\boldsymbol{\Sigma}_{yx}=\boldsymbol{\Sigma}_{xy}^{\mathrm{T}}=\boldsymbol{G}\boldsymbol{\Sigma}_{xx}$

目录

滤波器基本原理

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

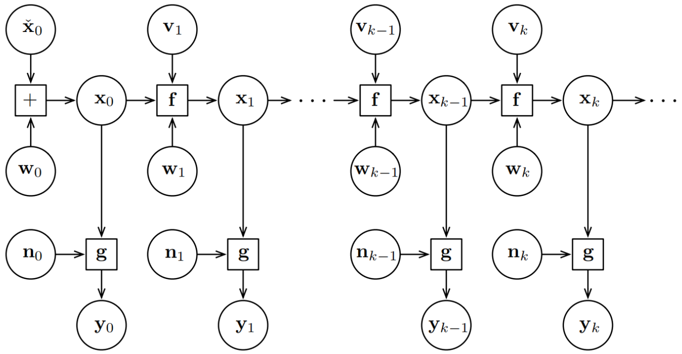
$\check{\boldsymbol{x}}_0$ 表示的是状态初始值

$\boldsymbol{v}_{1:k}$ 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，

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

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

滤波器基本原理

2. 贝叶斯滤波

根据贝叶斯公式，k时刻后验概率密度可以表示为

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

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，

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

滤波器基本原理

2. 贝叶斯滤波

经过以上化简，最终后验概率可以写为

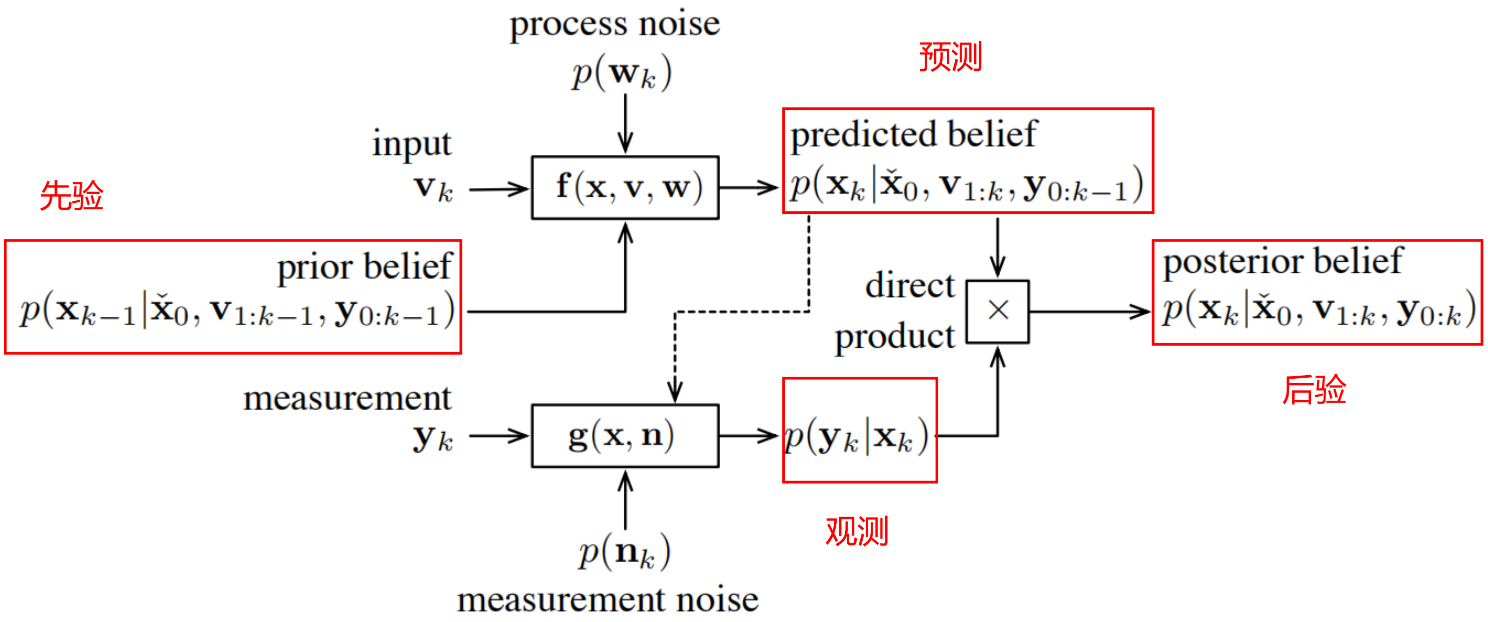
$$\underbrace{p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{0:k}\right)}_{\text{estimate}} = \eta \boxed{\underbrace{p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k\right)}_{\boldsymbol{g}(\cdot)}} \int \boxed{\underbrace{p\left(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1},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

观测　　　　预测　　　　先验

滤波器基本原理

2. 贝叶斯滤波

根据以上结果，可以画出贝叶斯滤波的信息流图如下



滤波器基本原理

2. 贝叶斯滤波

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. 卡尔曼滤波(KF)推导

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)

运动方程： $\boldsymbol{x}_k=\boldsymbol{F}\left(\boldsymbol{x}_{k-1},\boldsymbol{v}_k\right)+\boldsymbol{B}_{k-1}\boldsymbol{w}_k$

观测方程： $\boldsymbol{y}_k=\boldsymbol{G}\left(\boldsymbol{x}_k\right)+\boldsymbol{C}_k\boldsymbol{n}_k$

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。

滤波器基本原理

3. 卡尔曼滤波(KF)推导

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

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

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

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

3. 卡尔曼滤波(KF)推导

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。

4. 扩展卡尔曼滤波(EKF)推导

当运动方程或观测方程为非线性的时候，无法再利用之前所述的线性变化关系进行推导，常用的解决方法是进行线性化，把非线性方程一阶泰勒展开成线性。即

运动方程: $\boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{v}_k, \boldsymbol{w}_k\right) \approx \check{\boldsymbol{x}}_k + \boldsymbol{F}_{k-1}\left(\boldsymbol{x}_{k-1} - \hat{\boldsymbol{x}}_{k-1}\right) + \boldsymbol{B}_{k-1}\boldsymbol{w}_k$

观测方程： $\boldsymbol{y}_k = \boldsymbol{g}\left(\boldsymbol{x}_k, \boldsymbol{n}_k\right) \approx \check{\boldsymbol{y}}_k + \boldsymbol{G}_k\left(\boldsymbol{x}_k - \check{\boldsymbol{x}}_k\right) + \boldsymbol{C}_k\boldsymbol{n}_k$

其中

$$\check{\boldsymbol{x}}_k = \boldsymbol{f}\left(\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{v}_k, \boldsymbol{0}\right)$$

$$\check{\boldsymbol{y}}_k = \boldsymbol{g}\left(\check{\boldsymbol{x}}_k, \boldsymbol{0}\right)$$

$$\boldsymbol{F}_{k-1} = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{v}_k, \boldsymbol{w}_k)}{\partial \boldsymbol{x}_{k-1}}\right|_{\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{v}_k, \boldsymbol{0}}$$

$$\boldsymbol{G}_k = \left.\frac{\partial \boldsymbol{g}(\boldsymbol{x}_k, \boldsymbol{n}_k)}{\partial \boldsymbol{x}_k}\right|_{\check{\boldsymbol{x}}_k, \boldsymbol{0}}$$

$$\boldsymbol{B}_{k-1} = \left.\frac{\partial \boldsymbol{f}(\boldsymbol{x}_{k-1}, \boldsymbol{v}_k, \boldsymbol{w}_k)}{\partial \boldsymbol{w}_k}\right|_{\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{v}_k, \boldsymbol{0}}$$

$$\boldsymbol{C}_k = \left.\frac{\partial \boldsymbol{g}(\boldsymbol{x}_k, \boldsymbol{n}_k)}{\partial \boldsymbol{n}_k}\right|_{\check{\boldsymbol{x}}_k, \boldsymbol{0}}$$

滤波器基本原理

4. 扩展卡尔曼滤波(EKF)推导

根据该线性化展开结果，可以得到**预测状态**的统计学特征为

$$E[\boldsymbol{x}_k] \approx \check{\boldsymbol{x}}_k + \boldsymbol{F}_{k-1}(\boldsymbol{x}_{k-1} - \hat{\boldsymbol{x}}_{k-1}) + \underbrace{E[\boldsymbol{B}_{k-1}\boldsymbol{w}_k]}_{0}$$

$$E\left[(\boldsymbol{x}_k - E[\boldsymbol{x}_k])(\boldsymbol{x}_k - E[\boldsymbol{x}_k])^{\mathrm{T}}\right] \approx \underbrace{E\left[\boldsymbol{B}_{k-1}\boldsymbol{w}_k\boldsymbol{w}_k^{\mathrm{T}}\boldsymbol{B}_{k-1}^{\mathrm{T}}\right]}_{\boldsymbol{B}_{k-1}\boldsymbol{Q}_k\boldsymbol{B}_{k-1}^{\mathrm{T}}}$$

即 $$p(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}, \boldsymbol{v}_k) \approx \mathcal{N}\left(\check{\boldsymbol{x}}_k + \boldsymbol{F}_{k-1}(\boldsymbol{x}_{k-1} - \hat{\boldsymbol{x}}_{k-1}), \boldsymbol{B}_{k-1}\boldsymbol{Q}_k\boldsymbol{B}_{k-1}^{\mathrm{T}}\right)$$

同理，可得到**观测**的统计学特征为

$$E[\boldsymbol{y}_k] \approx \check{\boldsymbol{y}}_k + \boldsymbol{G}_k(\boldsymbol{x}_k - \check{\boldsymbol{x}}_k) + \underbrace{E[\boldsymbol{C}_k\boldsymbol{n}_k]}_{0}$$

$$E\left[(\boldsymbol{y}_k - E[\boldsymbol{y}_k])(\boldsymbol{y}_k - E[\boldsymbol{y}_k])^{\mathrm{T}}\right] \approx \underbrace{E\left[\boldsymbol{C}_k\boldsymbol{n}_k\boldsymbol{n}_k^{\mathrm{T}}\boldsymbol{C}_k^{\mathrm{T}}\right]}_{\boldsymbol{C}_k\boldsymbol{R}_k\boldsymbol{C}_k^{\mathrm{T}}}$$

即 $$p(\boldsymbol{y}_k \mid \boldsymbol{x}_k) \approx \mathcal{N}\left(\check{\boldsymbol{y}}_k + \boldsymbol{G}_k(\boldsymbol{x}_k - \check{\boldsymbol{x}}_k), \boldsymbol{C}_k\boldsymbol{R}_k\boldsymbol{C}_k^{\mathrm{T}}\right)$$

4. 扩展卡尔曼滤波(EKF)推导

后面即可按照线性高斯下推导卡尔曼滤波的步骤进行推导，最终得到经典五个公式

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

$$\check{\boldsymbol{x}}_k = \boldsymbol{f}\left(\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{v}_k, \mathbf{0}\right)$$

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

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

$$\hat{\boldsymbol{x}}_k = \check{\boldsymbol{x}}_k + \boldsymbol{K}_k\left(\boldsymbol{y}_k - \boldsymbol{g}\left(\check{\boldsymbol{x}}_k, \mathbf{0}\right)\right)$$

5. 迭代扩展卡尔曼滤波(IEKF)推导

由于非线性模型中做了线性化近似，当非线性程度越强时，误差就会较大，但是由于线性化的工作点离真值越近，线性化的误差就越小，因此解决该问题的一个方法是，通过迭代逐渐找到准确的线性化点，从而提高精度。

在EKF的推导中，其他保持不变，仅改变观测的线性化工作点，则有

$$\boldsymbol{g}\left(\boldsymbol{x}_k, \boldsymbol{n}_k\right) \approx \boldsymbol{y}_{\mathrm{op}, k} + \boldsymbol{G}_k\left(\boldsymbol{x}_k - \boldsymbol{x}_{\mathrm{op}, k}\right) + \boldsymbol{C}_k \boldsymbol{n}_k$$

其中

$$\boldsymbol{y}_{\mathrm{op}, k} = \boldsymbol{g}\left(\boldsymbol{x}_{\mathrm{op}, k}, \mathbf{0}\right)$$

$$\boldsymbol{G}_k = \left.\frac{\partial \boldsymbol{g}(\boldsymbol{x}_k, \boldsymbol{n}_k)}{\partial \boldsymbol{x}_k}\right|_{\boldsymbol{x}_{\mathrm{op}, \boldsymbol{k}}, \mathbf{0}}$$

$$\boldsymbol{C}_k = \left.\frac{\partial \boldsymbol{g}(\boldsymbol{x}_k, \boldsymbol{n}_k)}{\partial \boldsymbol{n}_k}\right|_{\boldsymbol{x}_{\mathrm{op}, \boldsymbol{k}}, \mathbf{0}}$$

5. 迭代扩展卡尔曼滤波(IEKF)推导

按照与之前同样的方式进行推导，可得到滤波的校正过程为

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

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

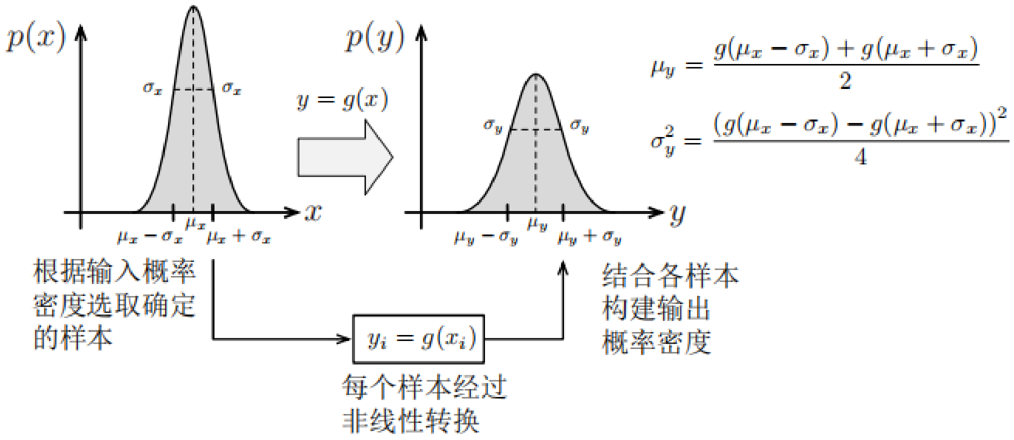
$$\hat{\boldsymbol{x}}_k = \check{\boldsymbol{x}}_k + \boldsymbol{K}_k \left(\boldsymbol{y}_k - \boldsymbol{y}_{\mathrm{op},k} - \boldsymbol{G}_k \left(\check{\boldsymbol{x}}_k - \boldsymbol{x}_{\mathrm{op},k} \right) \right)$$

可见唯一的区别是后验均值 $\hat{\boldsymbol{x}}_k$ 更新的公式与之前有所不同。

滤波过程中，反复执行上面3个公式，以上次的后验均值作为本次的线性化工作点，即可达到减小非线性误差的目的。

6. sigmapoint卡尔曼滤波(UKF/SPKF)推导

该方法的核心思想是，通过采样一部分sigmapoint点，传入非线性函数，通过计算这些点的分布，来近似概率密度函数。



滤波器基本原理

1)sigmapoint变换原理

a. 根据输入概率密度 $\mathcal{N}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_{xx})$ 计算出$2L+1$个sigmapoint，即

$$\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{\Sigma}_{xx}$$ (Cholesky分解，$\boldsymbol{L}$ 为下三角矩阵)

$$\boldsymbol{x}_0 = \boldsymbol{\mu}_x$$

$$\boldsymbol{x}_i = \boldsymbol{\mu}_x + \sqrt{L+\kappa}\,\mathrm{col}_i\,\boldsymbol{L}$$

$$\boldsymbol{x}_{i+L} = \boldsymbol{\mu}_x - \sqrt{L+\kappa}\,\mathrm{col}_i\,\boldsymbol{L}$$

其中

$$\boldsymbol{\mu}_x = \sum_{i=0}^{2L} \alpha_i \boldsymbol{x}_i$$

$$\boldsymbol{\Sigma}_{xx} = \sum_{i=0}^{2L} \alpha_i (\boldsymbol{x}_i - \boldsymbol{\mu}_x)(\boldsymbol{x}_i - \boldsymbol{\mu}_x)^{\mathrm{T}}$$

$$\alpha_i = \begin{cases} \frac{\kappa}{L+\kappa} & i=0 \\ \frac{1}{2}\frac{1}{L+\kappa} & \text{others} \end{cases}$$

b. 把每个sigmapoint单独带入非线性函数 g(·) 中；

$$\boldsymbol{y}_i = \boldsymbol{g}(\boldsymbol{x}_i), \quad i = 0 \ldots 2L$$

c. 输出概率的均值通过下面的式子计算；

$$\boldsymbol{\mu}_y = \sum_{i=0}^{2L} \alpha_i \boldsymbol{y}_i$$

d. 输出概率的协方差通过下面的式子计算。

$$\boldsymbol{\Sigma}_{yy} = \sum_{i=0}^{2L} \alpha_i (\boldsymbol{y}_i - \boldsymbol{\mu}_y)(\boldsymbol{y}_i - \boldsymbol{\mu}_y)^{\mathrm{T}}$$

滤波器基本原理

2) 预测步骤

a. 把先验置信度和运动噪声按如下方式堆叠到一起

$$\boldsymbol{\mu}_z = \begin{bmatrix} \hat{\boldsymbol{x}}_{k-1} \\ \mathbf{0} \end{bmatrix}$$

$$\boldsymbol{\Sigma}_{zz} = \begin{bmatrix} \hat{\boldsymbol{P}}_{k-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Q}_k \end{bmatrix}$$

b. 将 $\{\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_{zz}\}$ 转化为sigmapoint表示

$$\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{\Sigma}_{zz}$$

$$\boldsymbol{z}_0 = \boldsymbol{\mu}_z$$

$$\boldsymbol{z}_i = \boldsymbol{\mu}_z + \sqrt{L+\kappa}\,\mathrm{col}_i\,\boldsymbol{L}$$

$$\boldsymbol{z}_{i+L} = \boldsymbol{\mu}_z - \sqrt{L+\kappa}\,\mathrm{col}_i\,\boldsymbol{L}$$

c. 对每个 sigmapoint 展开为状态和运动噪声的形式

$$\boldsymbol{z}_i = \begin{bmatrix} \hat{\boldsymbol{x}}_{k-1,i} \\ \boldsymbol{w}_{k,i} \end{bmatrix}$$

接着将每个 sigmapoint 代入非线性运动模型进行精确求解

$$\check{\boldsymbol{x}}_{k,i} = \boldsymbol{f}\left(\hat{\boldsymbol{x}}_{k-1,i}, \boldsymbol{v}_k, \boldsymbol{w}_{k,i}\right), \quad i = 0 \ldots 2L$$

d. 将转换后的 sigmapoint 重新组合成预测置信度

$$\check{\boldsymbol{x}}_k = \sum_{i=0}^{2L} \alpha_i \check{\boldsymbol{x}}_{k,i}$$

$$\check{P}_k = \sum_{i=0}^{2L} \alpha_i \left(\check{\boldsymbol{x}}_{k,i} - \check{\boldsymbol{x}}_k\right)\left(\check{\boldsymbol{x}}_{k,i} - \check{\boldsymbol{x}}_k\right)^{\mathrm{T}}$$

其中

$$\alpha_i = \begin{cases} \frac{\kappa}{L+\kappa} & i = 0 \\ \frac{1}{2}\frac{1}{L+\kappa} & \text{others} \end{cases}$$

滤波器基本原理

3) 校正步骤

a. 把预测置信度和观测噪声按如下方式堆叠到一起

$$\boldsymbol{\mu}_z = \begin{bmatrix} \check{x}_k \\ \mathbf{0} \end{bmatrix} \qquad \boldsymbol{\Sigma}_{zz} = \begin{bmatrix} \check{\boldsymbol{P}}_k & \mathbf{0} \\ \mathbf{0} & \boldsymbol{R}_k \end{bmatrix}$$

b. 将 $\{\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_{zz}\}$ 转化为sigmapoint表示

$$\boldsymbol{L}\boldsymbol{L}^{\mathrm{T}} = \boldsymbol{\Sigma}_{zz}$$
$$\boldsymbol{z}_0 = \boldsymbol{\mu}_z$$
$$\boldsymbol{z}_i = \boldsymbol{\mu}_z + \sqrt{L+\kappa}\,\mathrm{col}_i\,\boldsymbol{L}$$
$$\boldsymbol{z}_{i+L} = \boldsymbol{\mu}_z - \sqrt{L+\kappa}\,\mathrm{col}_i\,\boldsymbol{L}$$

c. 对每个 sigmapoint 展开为状态和观测噪声的形式

$$\boldsymbol{z}_i = \begin{bmatrix} \check{\boldsymbol{x}}_{k,i} \\ \boldsymbol{n}_{k,i} \end{bmatrix}$$

将每个sigmapoint代入非线性观测模型进行精确求解

$$\check{\boldsymbol{y}}_{k,i} = \boldsymbol{g}\left(\check{\boldsymbol{x}}_{k,i}, \boldsymbol{n}_{k,i}\right), \quad i = 0 \ldots 2L$$

d. 将转换后的 sigmapoint 重新组合得到最终结果

$$\boldsymbol{\mu}_{y,k} = \sum_{i=0}^{2L} \alpha_i \check{\boldsymbol{y}}_{k,i}$$
$$\boldsymbol{\Sigma}_{yy,k} = \sum_{i=0}^{2L} \alpha_i \left(\check{\boldsymbol{y}}_{k,i} - \boldsymbol{\mu}_{y,k}\right)\left(\check{\boldsymbol{y}}_{k,i} - \boldsymbol{\mu}_{y,k}\right)^{\mathrm{T}}$$
$$\boldsymbol{\Sigma}_{xy,k} = \sum_{i=0}^{2L} \alpha_i \left(\check{\boldsymbol{x}}_{k,i} - \check{\boldsymbol{x}}_k\right)\left(\check{\boldsymbol{y}}_{k,i} - \boldsymbol{\mu}_{y,k}\right)^{\mathrm{T}}$$

其中

$$\alpha_i = \begin{cases} \frac{\kappa}{L+\kappa} & i = 0 \\ \frac{1}{2}\frac{1}{L+\kappa} & \text{others} \end{cases}$$

将这些式子带入到上面广义高斯滤波的校正步骤方程中，便完成了最终的校正步骤。

滤波器基本原理

7. 粒子滤波(PF)推导

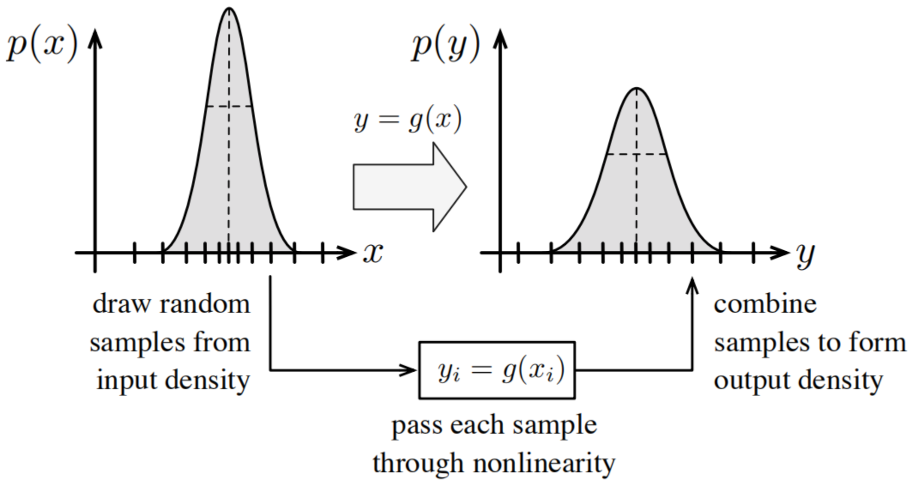
1) 蒙特卡罗方法原理

蒙特卡罗本质是基于大数定律，主要包括3步：

1) 根据输入的概率密度采集样本；

2) 通过非线性函数对每个样本进行转换；

3) 从转换的样本中重新计算概率密度。



蒙特卡罗方法原理示意图

7. 粒子滤波(PF)推导

2) 粒子滤波原理

a. 从由先验和运动噪声的联合概率密度中抽取 M 个样本

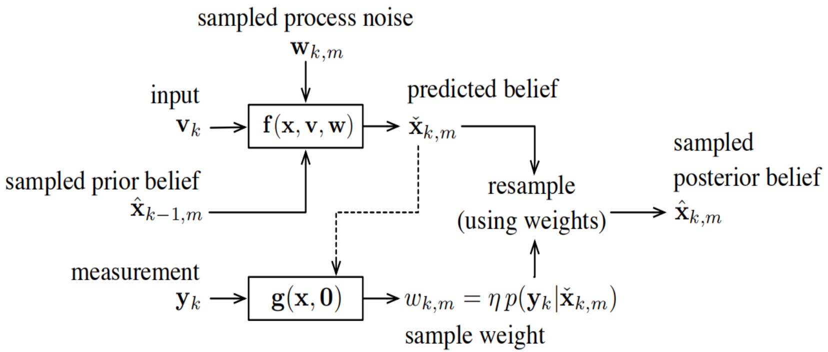
$$\begin{bmatrix} \hat{\boldsymbol{x}}_{k-1,m} \\ \boldsymbol{w}_{k,m} \end{bmatrix} \leftarrow p\left(\boldsymbol{x}_{k-1} \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k-1}, \boldsymbol{y}_{1:k-1}\right) p\left(\boldsymbol{w}_k\right)$$

其中，m为唯一的粒子序号。

b. 使用V k 得到后验概率分布函数的预测，可以将每个先验粒子和噪声样本代入非线性运动模型，

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

这些新的“预测粒子”共同近似刻画了概率密度 $p\left(\boldsymbol{x}_k \mid \check{\boldsymbol{x}}_0, \boldsymbol{v}_{1:k}, \boldsymbol{y}_{1:k-1}\right)$。



8. 粒子滤波(PF)推导

2) 粒子滤波原理

c. 结合 y_k 对后验概率进行校正，主要分两步：

第1步，根据每个粒子的期望后验和预测后验的收敛程度，对每个粒子赋予权值；

$$w_{k,m}=\eta p\left(\boldsymbol{y}_k \mid \check{\boldsymbol{x}}_{k,m}\right)$$

其中 η 为归一化系数。
在实践中，通常使用非线性观测模型来模拟期望的传感器读数，

$$\check{\boldsymbol{y}}_{k,m}=\boldsymbol{g}\left(\check{\boldsymbol{x}}_{k,m}, \mathbf{0}\right)$$

假设 $p\left(\boldsymbol{y}_k \mid \check{\boldsymbol{x}}_{k,m}\right)=p\left(\boldsymbol{y}_k \mid \check{\boldsymbol{y}}_{k,m}\right)$，其中等式右边的概率已知。

第2步，根据赋予的权重，对每个粒子进行重要性重采样

$$\hat{\boldsymbol{x}}_{k,m} \xleftarrow{\text{重要性重采样}} \left\{\check{\boldsymbol{x}}_{k,m}, w_{k,m}\right\}$$

8. PF、EKF、UKF对比

举例对比三种方法的不同：以一维非线性函数 $f(x)=x^2$ 为例，其概率密度为 $\mathcal{N}\left(\mu_x \sigma_x^2\right)$

1) 蒙特卡罗方法

根据输入密度获得的任意样本由下式给出

$$x_i=\mu_x+\delta x_i, \quad \delta x_i \leftarrow \mathcal{N}\left(0, \sigma_x^2\right)$$

将样本代入非线性函数中，可以得到

$$\begin{aligned} y_i &= f\left(x_i\right) \\ &= f\left(\mu_x+\delta x_i\right) \\ &= \left(\mu_x+\delta x_i\right)^2 \\ &= \mu_x^2+2 \mu_x \delta x_i+\delta x_i^2 \end{aligned}$$

对式子两边求期望，可以得到输出的均值

$$\begin{aligned} \mu_y &= E\left[y_i\right] \\ &= \mu_x^2+2 \mu_x \underbrace{E\left[\delta x_i\right]}_{0}+\underbrace{E\left[\delta x_i^2\right]}_{\sigma_x^2} \\ &= \mu_x^2+\sigma_x^2 \end{aligned}$$

同样的，我们可以推导出输出的方差

$$\begin{aligned} \sigma_y^2 &= E\left[\left(y_i-\mu_y\right)^2\right] \\ &= E\left[\left(2 \mu_x \delta x_i+\delta x_i^2-\sigma_x^2\right)^2\right] \\ &= \underbrace{E\left[\delta x_i^4\right]}_{3 \sigma_x^4}+4 \mu_x \underbrace{E\left[\delta x_i^3\right]}_{0}+\left(4 \mu_x^2-2 \sigma_x^2\right) \underbrace{E\left[\delta x_i^2\right]}_{\sigma_x^2}-4 \mu_x \sigma_x^2 \underbrace{E\left[\delta x_i\right]}_{0}+\sigma_x^4 \\ &= 4 \mu_x^2 \sigma_x^2+2 \sigma_x^4 \end{aligned}$$

即，使用蒙特卡罗方法与无数个样本，闭式地求得后验概率的前两阶矩

8. PF、EKF、UKF对比

2) 线性化方法

对非线性函数在输入概率密度的均值处进行线性化，可得

$$
\begin{aligned}
y_i &= f\left(\mu_x + \delta x_i\right) \\
&\approx \underbrace{f\left(\mu_x\right)}_{\mu_x^2} + \underbrace{\left.\frac{\partial f}{\partial x}\right|_{\mu_x}}_{2\mu_x} \delta x_i \\
&= \mu_x^2 + 2\mu_x \delta x_i
\end{aligned}
$$

对等式两边取期望，得到

$$
\mu_y = E\left[y_i\right] = \mu_x^2 + 2\mu_x \underbrace{E\left[\delta x_i\right]}_{0} = \mu_x^2
$$

对于输出的方差，有

$$
\begin{aligned}
\sigma_y^2 &= E\left[\left(y_i - \mu_y\right)^2\right] \\
&= E\left[\left(2\mu_x \delta x_i\right)^2\right] \\
&= 4\mu_x^2 \sigma_x^2
\end{aligned}
$$

8. PF、EKF、UKF对比

3) sigmapoint变换

在维度$L=1$的情况下，有$2L+1=3$个 sigmapoint

$$
\begin{aligned}
x_0 &= \mu_x \\
x_1 &= \mu_x + \sqrt{1+\kappa}\sigma_x \\
x_2 &= \mu_x - \sqrt{1+\kappa}\sigma_x
\end{aligned}
$$

将每个 sigmapoint 代入非线性函数中，得

$$
y_0 = f\left(x_0\right) = \mu_x^2
$$

$$
\begin{aligned}
y_1 &= f\left(x_1\right) \\
&= \left(\mu_x + \sqrt{1+\kappa}\sigma_x\right)^2 \\
&= \mu_x^2 + 2\mu_x\sqrt{1+\kappa}\sigma_x + (1+\kappa)\sigma_x^2
\end{aligned}
$$

$$
\begin{aligned}
y_2 &= f\left(x_2\right) \\
&= \left(\mu_x - \sqrt{1+\kappa}\sigma_x\right)^2 \\
&= \mu_x^2 - 2\mu_x\sqrt{1+\kappa}\sigma_x + (1+\kappa)\sigma_x^2
\end{aligned}
$$

8. PF、EKF、UKF对比

3) sigmapoint变换

输出的均值为

$$
\begin{aligned}
\mu_y =& \frac{1}{1+\kappa}\left(\kappa y_0 + \frac{1}{2}\sum_{i=1}^{2} y_i\right) \\
=& \frac{1}{1+\kappa}\left(\kappa \mu_x^2 + \frac{1}{2}\left(\mu_x^2 + 2\mu_x\sqrt{1+\kappa}\sigma_x + (1+\kappa)\sigma_x^2 + \mu_x^2 - 2\mu_x\sqrt{1+\kappa}\sigma_x + (1+\kappa)\sigma_x^2\right)\right) \\
=& \frac{1}{1+\kappa}\left(\kappa\mu_x^2 + \mu_x^2 + (1+\kappa)\sigma_x^2\right) \\
=& \mu_x^2 + \sigma_x^2
\end{aligned}
$$

输出的方差为

$$
\begin{aligned}
\sigma_y^2 =& \frac{1}{1+\kappa}\left(\kappa\left(y_0-\mu_y\right)^2 + \frac{1}{2}\sum_{i=1}^{2}\left(y_i-\mu_y\right)^2\right) \\
=& \frac{1}{1+\kappa}\left(\kappa\sigma_x^4 + \frac{1}{2}\left(\left(2\mu_x\sqrt{1+\kappa}\sigma_x + \kappa\sigma_x^2\right)^2 \quad + \left(-2\mu_x\sqrt{1+\kappa}\sigma_x + \kappa\sigma_x^2\right)^2\right)\right) \\
=& \frac{1}{1+\kappa}\left(\kappa\sigma_x^4 + 4(1+\kappa)\mu_x^2\sigma_x^2 + \kappa^2\sigma_x^4\right) \\
=& 4\mu_x^2\sigma_x^2 + \kappa\sigma_x^4
\end{aligned}
$$

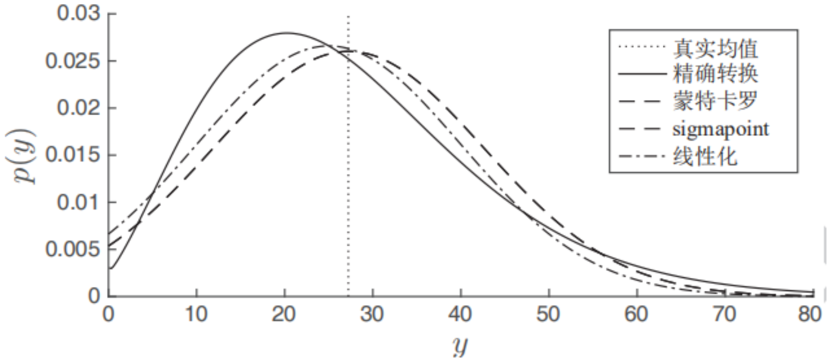
8. PF、EKF、UKF对比

4) 总结

蒙特卡罗： $\mu_y = \mu_x^2 + \sigma_x^2$ $\sigma_y^2 = 4\mu_x^2\sigma_x^2 + 2\sigma_x^4$

线 性 化： $\mu_y = \mu_x^2$ $\sigma_y^2 = 4\mu_x^2\sigma_x^2$

sigmapoint： $\mu_y = \mu_x^2 + \sigma_x^2$ $\sigma_y^2 = 4\mu_x^2\sigma_x^2 + \kappa\sigma_x^4$



a. PF是精确度最高的，因为是无损的；

b. UKF可以通过调整k的值，达到与PF一样的效果，而且效率更高；

c. 线性化方法是有偏的，但是效率最高；

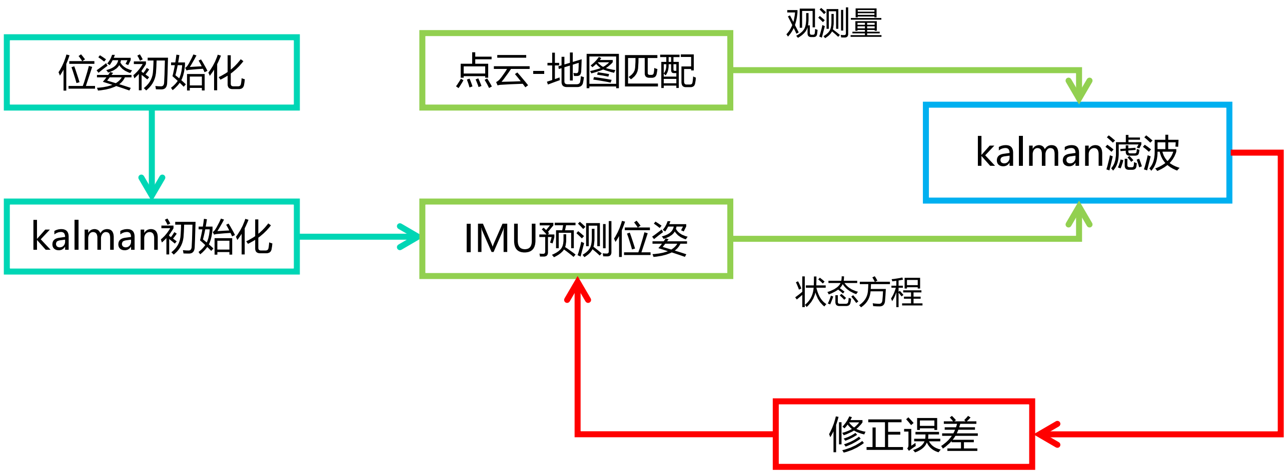
注意：实际应用中，应以模型的非线性程度和噪声分布来选择方法，不严重降低精度的情况下，选择效率和可靠性最高的。

目录

基于滤波器的融合

通过以上推导，滤波问题可以简单理解为“预测 + 观测 = 融合结果”。
结合实际点云地图中定位的例子，预测由IMU给出，观测即为激光雷达点云和地图匹配得到的姿态和位置。

融合流程用框图可以表示如下：



基于滤波器的融合

1. 基于误差状态的滤波

1) 状态方程

状态方程由误差方程得来，第三章已经完成误差方程的推导

$$\dot{\phi} = \phi \times \omega_{ie}^n - \delta\omega_{ib}^n$$

$$\delta\dot{V} = f^n \times \phi + \delta f^n$$

$$\delta\dot{P} = \delta V$$

为了写出完整的状态方程，需要把它们展开。

a) 姿态误差方程

由于

$$\phi = \begin{bmatrix}\phi_E & \phi_N & \phi_U\end{bmatrix}^T$$

$$\omega_{ie}^n = \begin{bmatrix}0 & \omega cosL & \omega sinL\end{bmatrix}^T$$

其中$\omega = 7.292115e(-5)rad/s$ 为地球自转角速度

而

$$\delta\omega_{ib}^n = C_b^n \begin{bmatrix}\varepsilon_x \\ \varepsilon_y \\ \varepsilon_z\end{bmatrix}$$

其中 ε 为陀螺仪的bias。

因此，姿态误差方程 $\dot{\phi} = \phi \times \omega_{ie}^n - \delta\omega_{ib}^n$ 可重新写为

$$\begin{bmatrix}\dot{\phi}_E \\ \dot{\phi}_N \\ \dot{\phi}_U\end{bmatrix} = \begin{bmatrix}0 & -\phi_U & \phi_N \\ \phi_U & 0 & -\phi_E \\ -\phi_N & \phi_E & 0\end{bmatrix} \begin{bmatrix}0 \\ \omega cosL \\ \omega sinL\end{bmatrix} - \boldsymbol{C}_b^n \begin{bmatrix}\varepsilon_x \\ \varepsilon_y \\ \varepsilon_z\end{bmatrix}$$

基于滤波器的融合

1. 基于误差状态的滤波

1) 状态方程

b) 速度误差方程

由于

$$f^n = \begin{bmatrix} f_E & f_N & f_U \end{bmatrix}^T$$

$$\delta f^n = C_b^n \begin{bmatrix} \nabla_x \\ \nabla_y \\ \nabla_z \end{bmatrix}$$

其中 ∇ 为加速度计的bias。

因此，速度误差方程 $\delta\dot{V} = f^n \times \phi + \delta f^n$ 可重新写为

$$\begin{bmatrix} \delta\dot{V}_E \\ \delta\dot{V}_N \\ \delta\dot{V}_U \end{bmatrix} = \begin{bmatrix} 0 & -f_U & f_N \\ f_U & 0 & -f_E \\ -f_N & f_E & 0 \end{bmatrix} \begin{bmatrix} \phi_E \\ \phi_N \\ \phi_U \end{bmatrix} + \boldsymbol{C}_b^n \begin{bmatrix} \nabla_x \\ \nabla_y \\ \nabla_z \end{bmatrix}$$

c) 位置误差方程

位置误差方程 $\delta\dot{P} = \delta V$ 可直接展开

$$\begin{bmatrix} \delta\dot{P}_E \\ \delta\dot{P}_N \\ \delta\dot{P}_U \end{bmatrix} = \begin{bmatrix} \delta V_E \\ \delta V_N \\ \delta V_U \end{bmatrix}$$

基于滤波器的融合

1. 基于误差状态的滤波

1) 状态方程

在滤波器中，状态方程一般写为如下形式

$$\dot{X} = F_t X + B_t W$$

根据以上展开的误差方程，若令状态量

$$X = \begin{bmatrix} \delta P^T & \delta V^T & \phi^T & \varepsilon^T & \nabla^T \end{bmatrix}^T$$

其中

$$\delta P = \begin{bmatrix} \delta P_E & \delta P_N & \delta P_U \end{bmatrix}^T$$

$$\delta V = \begin{bmatrix} \delta V_E & \delta V_N & \delta V_U \end{bmatrix}^T$$

$$\phi = \begin{bmatrix} \phi_E & \phi_N & \phi_U \end{bmatrix}^T$$

$$\varepsilon = \begin{bmatrix} \varepsilon_x & \varepsilon_y & \varepsilon_z \end{bmatrix}^T$$

$$\nabla = \begin{bmatrix} \nabla_x & \nabla_y & \nabla_z \end{bmatrix}^T$$

则有

$$F_t = \begin{bmatrix} 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & F_{23} & 0_{3\times3} & C_b^n \\ 0_{3\times3} & 0_{3\times3} & F_{33} & -C_b^n & 0_{3\times3} \\ \multicolumn{5}{c}{0_{3\times15}} \\ \multicolumn{5}{c}{0_{3\times15}} \end{bmatrix}$$

其中

$$F_{23} = \begin{bmatrix} 0 & -f_U & f_N \\ f_U & 0 & -f_E \\ -f_N & f_E & 0 \end{bmatrix}$$

$$F_{33} = \begin{bmatrix} 0 & 0 & 0 \\ -\omega sinL & 0 & 0 \\ -\omega cosL & 0 & 0 \end{bmatrix}$$

基于滤波器的融合

1. 基于误差状态的滤波

1) 状态方程

在状态方程右侧，还包含W，它代表的是器件噪声，在惯导问题中

$$W = \begin{bmatrix} w_{gx} & w_{gy} & w_{gz} & w_{ax} & w_{ay} & w_{az} \end{bmatrix}^T$$

其中w_g为陀螺仪的噪声，w_a为加速度计的噪声。

此时有

$$B_t = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & C_b^n \\ -C_b^n & 0_{3\times3} \\ 0_{6\times3} & 0_{6\times3} \end{bmatrix}$$

基于滤波器的融合

1. 基于误差状态的滤波

2) 观测方程

在滤波器中，观测方程一般写为

$$Y = G_t X + C_t N$$

此例中观测量有位置、失准角，则

$$Y = \begin{bmatrix} \delta P^T & \phi^T \end{bmatrix}^T$$

由于

$$X = \begin{bmatrix} \delta P^T & \delta V^T & \phi^T & \varepsilon^T & \nabla^T \end{bmatrix}^T$$

因此有

$$G_t = \begin{bmatrix} I_{3\times3} & 0_{3\times3} & 0_{3\times3} & 0_{3\times6} \\ 0_{3\times3} & 0_{3\times3} & I_{3\times3} & 0_{3\times6} \end{bmatrix}$$

$$C_t = \begin{bmatrix} I_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & I_{3\times3} \end{bmatrix}$$

而此处N为观测噪声

$$N = \begin{bmatrix} n_{P_E} & n_{P_N} & n_{P_U} & n_{\phi_E} & n_{\phi_N} & n_{\phi_U} \end{bmatrix}^T$$

注意，此处得到观测值 ϕ 的计算过程为：

先计算误差矩阵，

$$\boldsymbol{C}_n^{n'} = \boldsymbol{C}_b^{n'} (\boldsymbol{C}_b^n)^T$$

其中 $\boldsymbol{C}_b^{n'}$ 为当前时刻位姿的预测值，而 $\boldsymbol{C}_b^n$ 为姿态矩阵的观测值（此处指点云匹配得到的结果）。

而又根据下式，便可以得到 ϕ 。

$$\boldsymbol{C}_n^{n'} \approx \boldsymbol{I} - (\boldsymbol{\phi} \times)$$

基于滤波器的融合

1. 基于误差状态的滤波

3) 构建滤波器

构建滤波器，即把融合系统的状态方程和观测方程应用到kalman滤波的五个公式中。

注：上面推导的方程是连续时间的，要应用于离散时间，需要按照采样时间对其进行离散化。

状态方程离散化，可以写为

$$X_k = F_{k-1}X_{k-1} + B_{k-1}W_k$$

其中

$$F_{k-1} = I_{15\times 15} + F_t(\hat{X}_{k-1}, v_k, 0)T$$

$$B_{k-1} = B_t(\hat{X}_{k-1}, v_k, 0)T$$

其中，T为kalman的滤波周期。

对于观测方程，不需要乘以滤波周期，可以直接写出

$$Y_k = G_kX_k + C_kN_k$$

将以上各变量，带入kalman滤波的五个方程，即

$$\check{X}_k = F_{k-1}\hat{X}_k + B_{k-1}W_k$$

$$\check{P}_k = F_{k-1}\hat{P}_{k-1}F_{k-1}^{\mathrm{T}} + B_{k-1}Q_kB_{k-1}^{T}$$

$$K_k = \check{P}_kG_k^{\mathrm{T}}\left(G_k\check{P}_kG_k^{\mathrm{T}} + C_kR_kC_k^{T}\right)^{-1}$$

$$\hat{P}_k = (I - K_kG_k)\check{P}_k$$

$$\hat{x}_k = \check{x}_k + K_k\left(Y_k - G_k\check{X}_k\right)$$

重复以上过程，便可以持续得到传感器融合的位姿输出。

基于滤波器的融合

2. 基于导航信息的滤波

有些情况下，滤波的状态量不是使用误差量，而是直接使用导航信息，此时状态量为

$$X = \begin{bmatrix} P^T & V^T & q^T & \varepsilon^T & \nabla^T \end{bmatrix}^T$$

1) 状态方程

导航信息的微分方程为

$$\dot{P} = V$$

$$\dot{V} = C_b^n(f^b + \nabla + \omega_a) + g^n$$

$$\dot{q} = \frac{1}{2} q \circ \left(w^b + \varepsilon + \omega_g\right)$$

其中，C_b^n 可以由四元数和旋转矩阵的关系得到。

$$C_b^n = \begin{bmatrix} C_{1,1} & C_{1,2} & C_{1,3} \\ C_{2,1} & C_{2,2} & C_{2,3} \\ C_{3,1} & C_{3,2} & C_{3,3} \end{bmatrix}$$

其中

$$C_{1,1} = q_0^2 + q_1^2 - q_2^2 - q_3^2$$
$$C_{1,2} = 2\left(q_1 q_2 + q_0 q_3\right)$$
$$C_{1,3} = 2\left(q_1 q_3 - q_0 q_2\right)$$
$$C_{2,1} = 2\left(q_1 q_2 - q_0 q_3\right)$$
$$C_{2,2} = q_0^2 - q_1^2 + q_2^2 - q_3^2$$
$$C_{2,3} = 2\left(q_2 q_3 + q_0 q_1\right)$$
$$C_{3,1} = 2\left(q_1 q_3 + q_0 q_2\right)$$
$$C_{3,2} = 2\left(q_2 q_3 - q_0 q_1\right)$$
$$C_{3,3} = q_0^2 - q_1^2 - q_2^2 + q_3^2$$

基于滤波器的融合

2. 基于导航信息的滤波

因此状态方程可以写成如下形式

$$\dot{X} = f(X, v, \omega, g^n)$$

它具有非线性，将其线性化可得

$$\dot{X} = F_t X + B_t W + \bar{f}(g^n)$$

此处

$$F_t = \begin{bmatrix} 0_{3\times3} & I_{3\times3} & 0_{3\times4} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & F_{Vq} & 0_{3\times3} & C_b^n \\ 0_{4\times3} & 0_{4\times3} & F_{qq} & F_{q\varepsilon} & 0_{4\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times4} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times4} & 0_{3\times3} & 0_{3\times3} \end{bmatrix}$$

其中 $F_{Vq} = \frac{\partial[C_b^n(f^b+\nabla+\omega_a)+g^n]}{\partial q}$

因此

$$F_{Vq} = \begin{bmatrix} F_{Vq0} & F_{Vq1} & F_{Vq2} & F_{Vq3} \\ -F_{Vq3} & -F_{Vq2} & F_{Vq1} & F_{Vq0} \\ F_{Vq2} & -F_{Vq3} & -F_{Vq0} & F_{Vq1} \end{bmatrix}$$

$$F_{Vq0} = 2\left(q_0\tilde{f}_x - q_3\tilde{f}_y + q_2\tilde{f}_z\right)$$

$$F_{Vq1} = 2\left(q_1\tilde{f}_x + q_2\tilde{f}_y + q_3\tilde{f}_z\right)$$

$$F_{Vq2} = 2\left(-q_2\tilde{f}_x + q_1\tilde{f}_y + q_0\tilde{f}_z\right)$$

$$F_{Vq3} = 2\left(-q_3\tilde{f}_x - q_0\tilde{f}_y + q_1\tilde{f}_z\right)$$

$\tilde{f} = f^b + \nabla + \omega_a$ 代表的是测量的加速度

基于滤波器的融合

2. 基于导航信息的滤波

由第三章所述四元数相乘可转矩阵相乘的结论可知

$$F_{qq} = \frac{1}{2}\begin{bmatrix} 0 & -\tilde{\omega}_x & -\tilde{\omega}_y & -\tilde{\omega}_z \\ \tilde{\omega}_x & 0 & \tilde{\omega}_z & -\tilde{\omega}_y \\ \tilde{\omega}_y & -\tilde{\omega}_z & 0 & \omega_x \\ \tilde{\omega}_z & \tilde{\omega}_y & -\tilde{\omega}_x & 0 \end{bmatrix}$$

其中 $\tilde{\omega} = \omega + \varepsilon + w_g$ 代表的是测量的角速度

把四元数相乘转成矩阵相乘以后，还可得

$$F_{q\varepsilon} = \frac{1}{2}\begin{bmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix}$$

噪声项与前面基于误差状态的kalman一致，即

$$W = \begin{bmatrix} w_{gx} & w_{gy} & w_{gz} & w_{ax} & w_{ay} & w_{az} \end{bmatrix}^T$$

因此状态方程函数 f 对噪声项求导可得

$$B_t = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & C_b^n \\ F_{q\varepsilon} & 0_{4\times3} \\ 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix}$$

而常数项保持不变

$$\bar{f}(g^n) = \begin{bmatrix} 0_{3\times1} \\ g^n \\ 0_{10\times1} \end{bmatrix}$$

基于滤波器的融合

2. 基于导航信息的滤波
2) 观测方程

以基于点云地图定位为例，观测为位置和姿态，即

$$Y = \begin{bmatrix} P^T & q^T \end{bmatrix}^T$$

因此观测方程如下

$$Y = G_t X + C_t N$$

其中

$$G_t = \begin{bmatrix} I_{3\times3} & 0_{3\times3} & 0_{3\times4} & 0_{3\times6} \\ 0_{4\times3} & 0_{4\times3} & I_{4\times4} & 0_{4\times6} \end{bmatrix}$$

由于观测量形式变化带来观测噪声形式变化，即

$$N = \begin{bmatrix} n_{P_E} & n_{P_N} & n_{P_U} & n_{q_0} & n_{q_1} & n_{q_2} & n_{q_3} \end{bmatrix}^T$$

因此有

$$C_t = I_{7\times7}$$

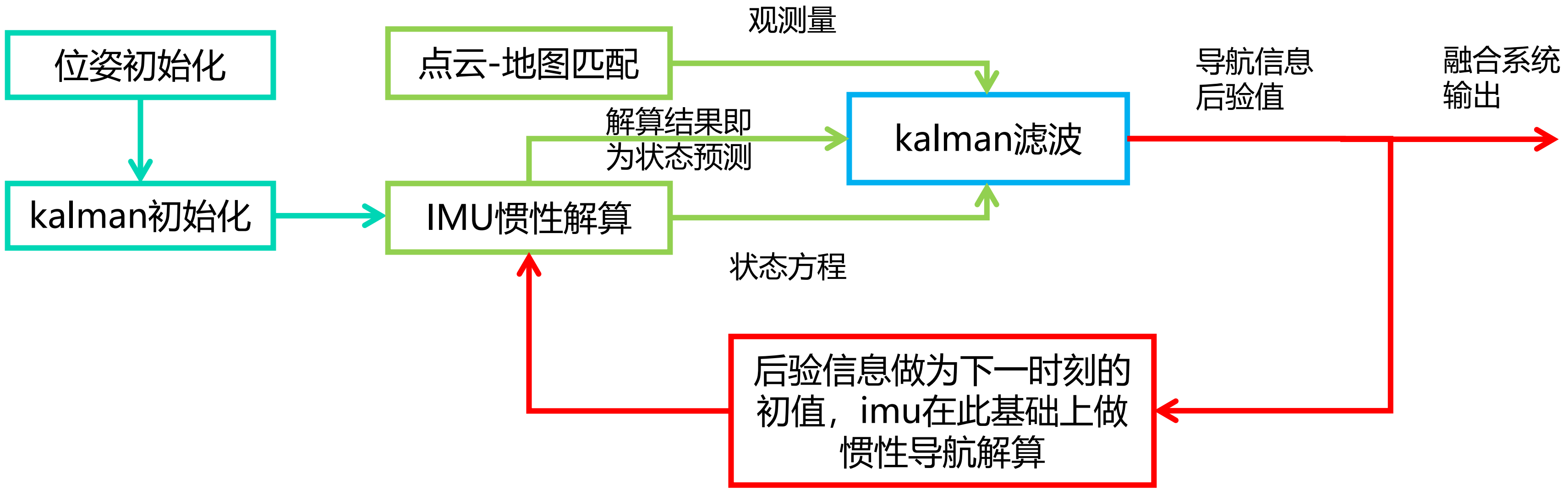
至此，已知基于导航信息方法下连续时间的线性化状态方程和观测方程，连续时间方程计算离散时间方程的方法刚已经介绍。得到离散时间方程后，便可以直接使用kalman进行预测(IMU)和观测(点云匹配)的融合。

基于滤波器的融合

3. kalman滤波实际使用流程

1) 基于导航信息的kalman滤波流程

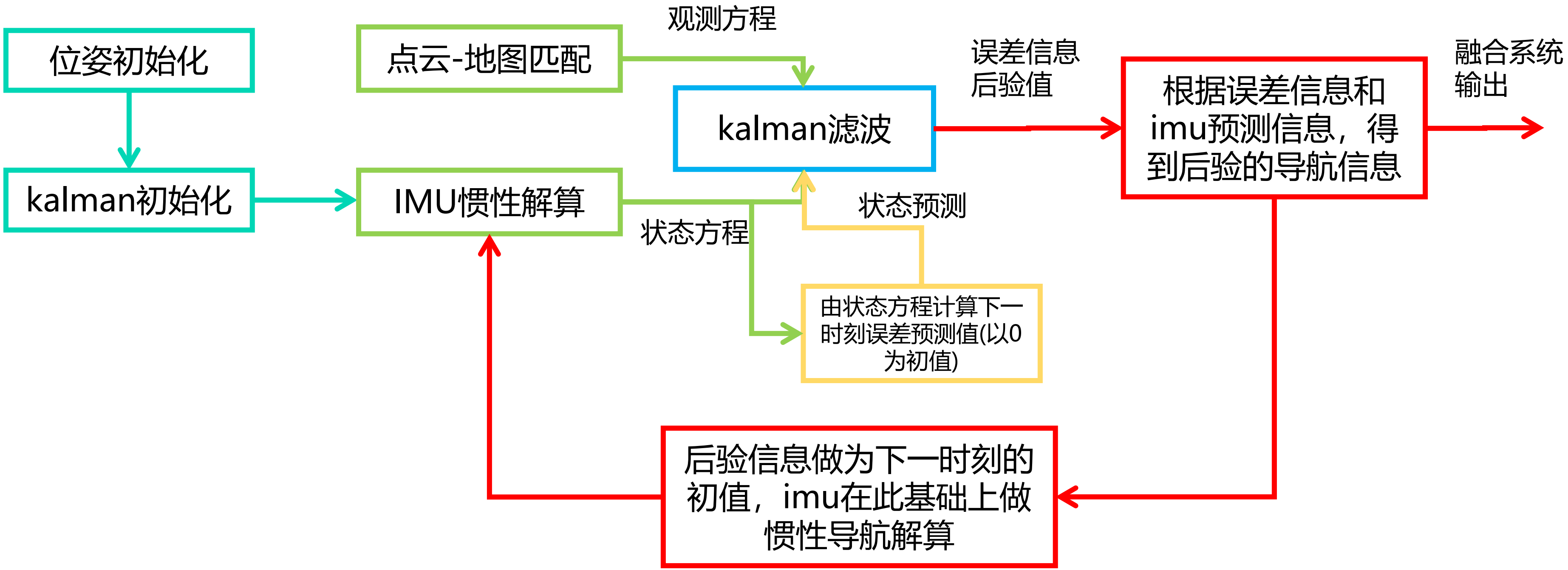
为了便于理解，先介绍基于导航信息的方法。



基于滤波器的融合

3. kalman滤波实际使用流程

2) 基于误差信息的kalman滤波流程



基于滤波器的融合

4. 一些讨论

1) 一般认为基于误差状态的kalman优于基于导航信息的kalman，理由是：

a. 误差状态自由度与实际位姿自由度相等，可避免过参数化；

b. 误差状态总是在0附近，因此线性化会更准确；

c. 误差状态是小量，因此二阶项可以直接忽略，方便雅可比计算；

d. 早期计算精度不高时，误差量相对于导航信息是小量，因此使用导航信息做状态时会被计算精度吃掉。

2) 惯性器件精度较高时，状态修正频率可以小于kalman滤波频率。

目录

观测性与观测度分析

1. 理解可观测性

以一个简单的方程组为例子

$$\begin{cases} x_1 + x_2 = 1 \\ x_1 + 2x_2 = 1 \end{cases}$$

显而易见，该方程组有解。
若令

$$X = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^T$$

$$Y = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$$

并分别以他们为状态量和观测量构建kalman方程，则该模型是”完全可观测的”。

相反，若有以下方程组

$$\begin{cases} x_1 + x_2 = 1 \\ 2x_1 + 2x_2 = 2 \end{cases}$$

若以它为基础构建kalman，则该模型是”不完全可观测的”。

一个模型的“可观测性”指的是能否由它的模型反推出状态。区分一个方程是否满足可观测的方法即看它的系数矩阵是否满秩。

观测性与观测度分析

2. 理解可观测度

仍以可观测的方程组为例，现在额外考虑了观测噪声，

$$\begin{cases} x_1 + x_2 = 1 + w_1 \\ x_1 + 2x_2 = 1 + w_2 \end{cases}$$

它的解为

$$\begin{cases} x_1 = 1 + 2w_1 - w_2 \\ x_2 = w_2 - w_1 \end{cases}$$

由于噪声是随机值，因此两个变量解的误差范围可以表示为

$$\begin{cases} \delta x_1 \leq |2w_1| + |w_2| \leq 3w \\ \delta x_2 \leq |w_2| + |w_1| \leq 2w \end{cases}$$

现在我们考虑另外一个方程组，

$$\begin{cases} x_1 + x_2 = 1 + w_1 \\ 2x_1 + x_2 = 1 + w_2 \end{cases}$$

按照同样的方式，可得到两个变量解得取值范围为

$$\begin{cases} \delta x_1 \leq 2w \\ \delta x_2 \leq 3w \end{cases}$$

可见，虽然两个方程都完全可观测，但变量解得精确范围却不同，即他们的“可观测度”不同。

观测性与观测度分析

2. 理解可观测度

判断可观测度的方法多种多样，其中一种常见的方法是SVD分解。

首先把方程写成如下形式

$$Y = AX$$

对系数矩阵进行SVD分解可得

$$A = U\Sigma V^T$$

其中 $\sum$ 为奇异值构成的对角阵，即

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

且有

$$\sigma_1 \geq \sigma_2$$

奇异值越大，变量的可观测度就越强。

按照这种方式，对上面两个例子进行SVD分解。

第一个例子，有

$$\Sigma = \begin{bmatrix} 2.618 & 0 \\ 0 & 0.382 \end{bmatrix}$$

但是对于第二个例子，同样有

$$\Sigma = \begin{bmatrix} 2.618 & 0 \\ 0 & 0.382 \end{bmatrix}$$

原因是，奇异值是按照大小顺序排列，并不是按照与状态量同样的顺序排列，因此我们只能知道在一个系统里，各变量的观测度差异有多少，但是却不知道那个状态量是精度更高的那个。

2. 理解可观测度

为了解决上面的问题，可以进行如下分析

首先把方程写成如下形式

$$X = A^{-1}Y = (U\Sigma V^T)^{-1}Y$$

由于奇异值分解中的U和V都是正交矩阵，因此

$$X = A^{-1}Y = V\Sigma^{-1}U^TY$$

进一步，还可以展开为

$$X = \sum_{i=1}^{2} \frac{u_i^T Y v_i}{\sigma_i}$$

其中有

$$U = \begin{bmatrix} u_1 & u_2 \end{bmatrix}$$
$$V = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$$

上式表示，X的解被分解为多个解向量的和，即

$$X = X_1 + X_2 = \begin{bmatrix} x_{11} \\ x_{22} \end{bmatrix} + \begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix}$$

带入具体的例子，比如例子1中，有

$$X_1 = \begin{bmatrix} 0.28 \\ 0.45 \end{bmatrix} \qquad X_2 = \begin{bmatrix} 0.72 \\ -0.45 \end{bmatrix}$$

把奇异值和状态量对应的方法是，看 X_i 中哪个状态量的绝对值最大，那么该状态量对应的奇异值就是 σ_i。比如，在X_1中有 0.45>0.28，那么状态量 x_2 对应的奇异值就是2.618，同样可知，状态量 x_1 对应的奇异值是0.382，因此 x_2的解精度更高，与之前的计算结果完全符合。

观测性与观测度分析

2. 理解可观测度

在第二个例子中，可以得到同样的结论。

$$X_1 = \begin{bmatrix} 0.45 \\ 0.28 \end{bmatrix} \qquad X_2 = \begin{bmatrix} -0.45 \\ 0.72 \end{bmatrix}$$

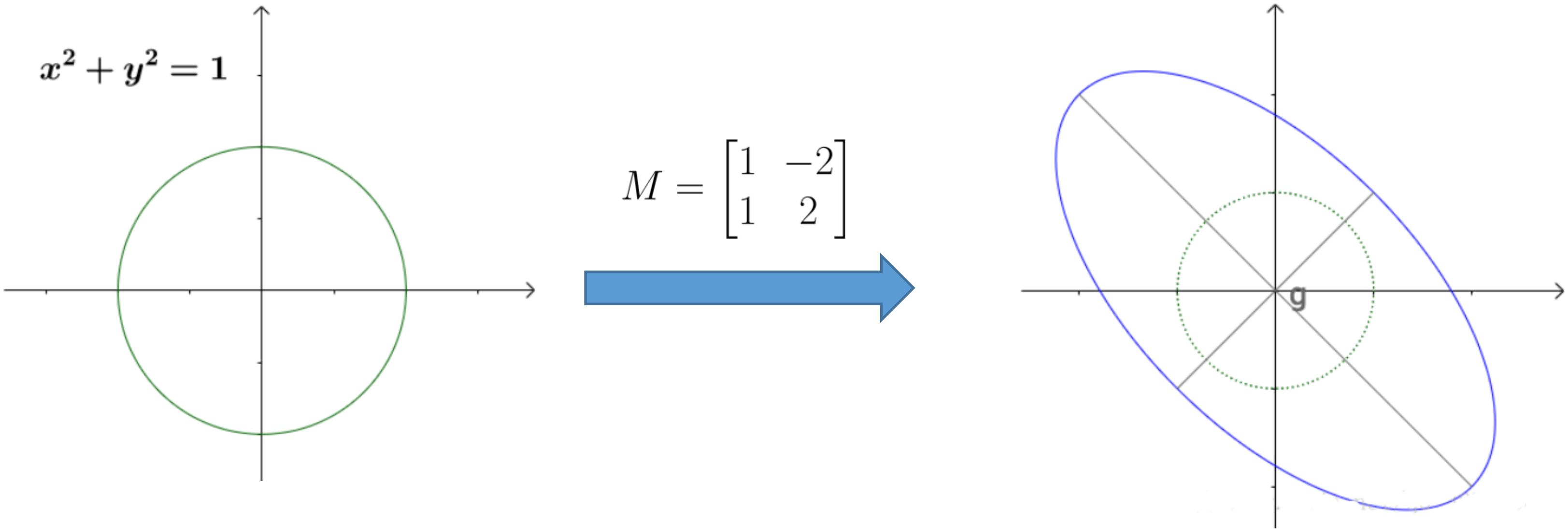
可以看出，状态量 x_1 对应的奇异值就是2.618，同样可知，状态量 x_2 对应的奇异值是0.382，因此 x_1 的解精度更高，与之前的计算结果也同样符合。

为了定量评价，此处的奇异值即可以作为可观测度。

观测性与观测度分析

3. 理解SVD

对于SVD为何能起到这种效果，需要做出进一步解释，为了便于理解，此处给出一种直观但不严谨的方式表述这个问题。



观测性与观测度分析

3. 理解SVD

对M进行奇异值分解

$$
\begin{aligned}
M &= \begin{bmatrix} 1 & -2 \\ 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} -0.707 & -0.707 \\ 0.707 & -0.707 \end{bmatrix} \begin{bmatrix} 2.828 & 0 \\ 0 & 1.414 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}
\end{aligned}
$$

M可以分解为两个部分

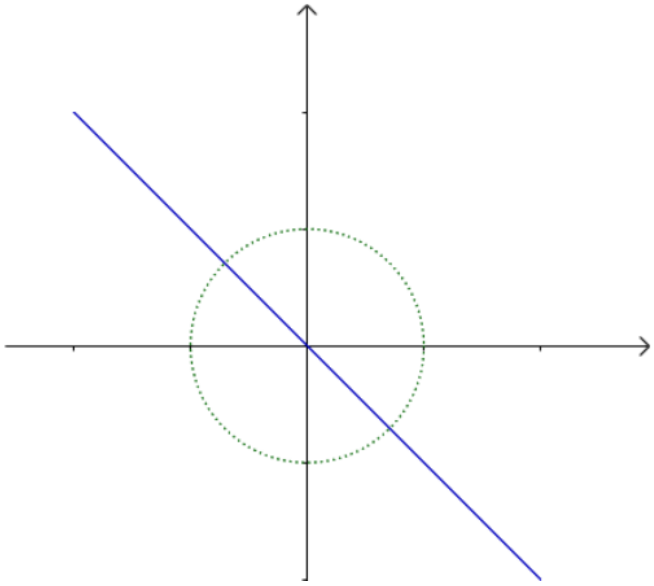
$$
\begin{aligned}
M = & \begin{bmatrix} -0.707 & -0.707 \\ 0.707 & -0.707 \end{bmatrix} \begin{bmatrix} 2.828 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + \\
& \begin{bmatrix} -0.707 & -0.707 \\ 0.707 & -0.707 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1.414 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}
\end{aligned}
$$

即把空间的变换拆分成了两次独立的变换，可以分开观察

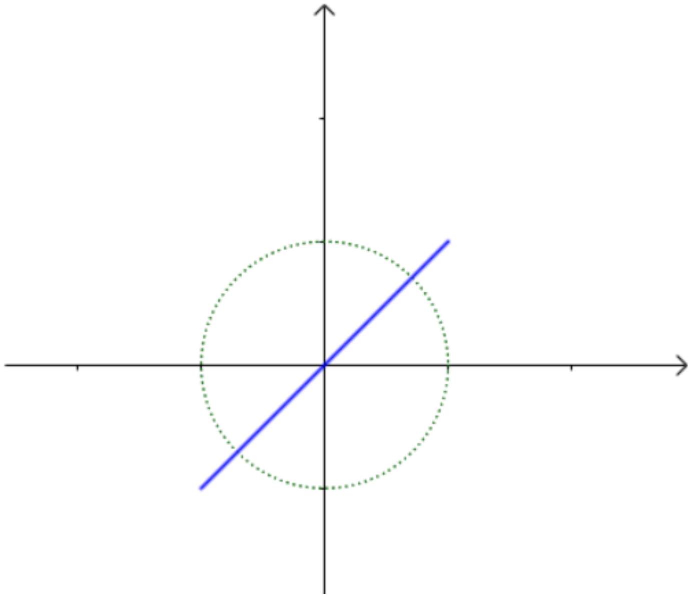
观测性与观测度分析

3. 理解SVD

第一个奇异值对应的变换



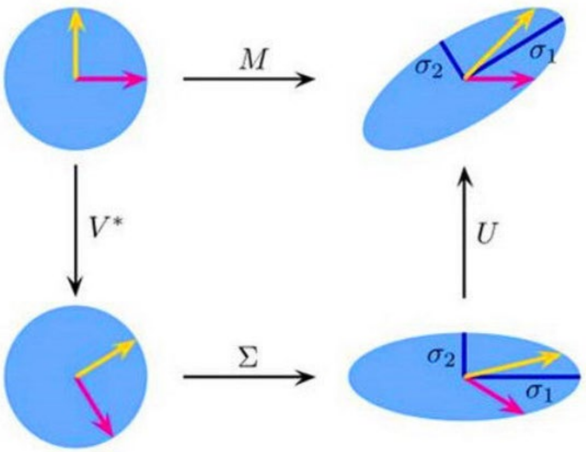
第二个奇异值对应的变换



观测性与观测度分析

3. 理解SVD

因此变换的整个流程可以概括为下图



$$M = U \cdot \Sigma \cdot V^*$$

物理含义：矩阵对向量的投影，可以分解成多个方向投影的和，这些方向上向量的分量，构成解空间。

在可观测度分析中的意义：观测在各方向上都有投影，按某个奇异值投影之后，哪那个变量对应的值最大，说明该奇异向量就是它的主方向，那么该奇异值就是它对应的奇异值了。

$$X = \sum_{i=1}^{n} \frac{u_i^T Y v_i}{\sigma_i}$$

观测性与观测度分析

4. 可观测性与可观测度的应用

1) 实际意义

可观测性：

在实际滤波模型中，当状态量较多或观测量较少时，会有部分状态量不可观测，这会导致滤波器不收敛，或收敛到一个错误的值上，并且对噪声的抗干扰性减弱，因此设计一个滤波模型后，需要对其进行可观测性分析。

可观测度：

当条件确实不能使所有状态量可观测时，应剔除不可观测的状态量，对模型进行裁剪。原因在于不可观测的状态量受噪声影响严重，并且会污染和它耦合的量。此外，对于多个均可观测的量，其可观测度不同，它们在kalman滤波中的收敛速度和收敛精度也会不同。

观测性与观测度分析

4. 可观测性与可观测度的应用

2) 定常系统可观测性分析

kalman滤波模型中的可观测性分析方法也由kalman本人发明，思路来源于他在现代控制理论中对系统模型进行的可观测性分析，二者形式上完全一致。

可观测性矩阵定义：

$$Q = \begin{bmatrix} G \\ GF \\ GF^2 \\ \cdots \\ GF^{n-1} \end{bmatrix}$$

其中 n 为滤波器维数。

Rudolf Kalman at his office in 1974.

观测性与观测度分析

4. 可观测性与可观测度的应用

2) 定常系统可观测性分析

可观测性矩阵的推导方法仍有多种，为了方便理解，此处仍选取解方程的思路来讲述。

在前述解方程的例子中，若能写成$Y = QX$ 的形式，则判断 Q是否满秩，即可判断系统是否完全可观测。

kalman滤波的模型为

$$\begin{cases} \dot{X} = FX \\ Y = GX \end{cases}$$

该式本质上是一个微分方程，因此可以直接使用微分方程的解，即有

$$Y = Ge^{Ft}X(0) = QX(0)$$

指数函数可以泰勒展开，即

$$e^{Ft} = e^{\bar{F}} = 1 + \bar{F} + \frac{\bar{F}^2}{2!} + \frac{\bar{F}^3}{3!} + \cdots$$

根据“卡莱-哈密顿定理”,因为 $\bar{F}$ 是n阶方阵，因此对于 $\forall k \geq n$，必有

$$\bar{F}^k = a_0 + a_1\bar{F} + a_2\bar{F}^2 + \cdots + a_{n-1}\bar{F}^{n-1}$$

因此，上面的指数函数必然能够写成下面的形式

$$e^{\bar{F}} = b_0 + b_1\bar{F} + b_2\bar{F}^2 + \cdots + b_{n-1}\bar{F}^{n-1}$$

即

$$Q = \sum_{i=0}^{n-1} b_i G F^i = \begin{bmatrix} b_0 I & b_1 I & \cdots & b_{n-1} I \end{bmatrix} \begin{bmatrix} G \\ G\bar{F} \\ \cdots \\ G\bar{F}^{n-1} \end{bmatrix}$$

观测性与观测度分析

4. 可观测性与可观测度的应用

2) 定常系统可观测性分析

因此，当下式成立时，系统满足可观测性

$$rank(Q) = rank(\begin{bmatrix} G \\ G\bar{F} \\ \cdots \\ G\bar{F}^{n-1} \end{bmatrix}) = n$$

3) 时变系统可观测性分析

在定位任务中，大部分系统的状态方程都是时变的，而上面的可观测性分析的推导只适合于定常系统。

此问题的解决办法是使用分段定常系统分析方法，即PWCS(Piece-Wise Constant Systems)，核心思路是把时变系统分成多个小的时间段，认为在每个时间段内是定常系统，随后找到整个系统观测性和各定常系统观测性之间的联系，即可解决问题。

具体证明过程细节参见文献：
Observability Analysis of Piece-Wise Constant Systems

观测性与观测度分析

4. 可观测性与可观测度的应用

3) 时变系统可观测性分析

首先，通过简单例子理解该思想的核心。

假设在0时刻有方程

$$\begin{cases} x_1 + x_2 = 1 \\ 2x_1 + 2x_2 = 2 \end{cases}$$

该方程显然不可解。

假设在1时刻有如下方程

$$\begin{cases} x_1 + 2x_2 = 2 \\ 2x_1 + 4x_2 = 4 \end{cases}$$

显然也不可解。

但是如果把两个方程联立，那么就可以解出各状态量。

PWCS的本质思想，就是看多个时刻的方程组联立之后，状态量是否可解。只不过在kalman标准模型下，有更严谨的表达形式。

观测性与观测度分析

4. 可观测性与可观测度的应用

3) 时变系统可观测性分析

a. 首先把系统划分成不同的时间段，每个时间段内对应一组状态方程和观测方程

$$\begin{cases} \dot{X}(t) = F_i X(t) \\ Y(t) = G_i X(t) \end{cases} \quad i = 1, 2, \ldots, r$$

第i时刻的可观测性矩阵可以表示为

$$Q_i = \begin{bmatrix} G_i \\ G_i F_i \\ G_i F_i^2 \\ \cdots \\ G_i F_i^{n-1} \end{bmatrix}$$

观测性与观测度分析

4. 可观测性与可观测度的应用

3) 时变系统可观测性分析

b. 根据不同时刻观测性矩阵，构建时变系统总的可观测性矩阵(TOM)

$$
Q_{TO}(r)=\begin{bmatrix} H_1 \\ H_1F_1 \\ \vdots \\ H_1F_1^{n-1} \\ -\,-\,-\,-\,-\,-\,-\,- \\ H_2F_1^{n-1} \\ H_2F_2F_1^{n-1} \\ \vdots \\ H_2F_2^{n-1}F_1^{n-1} \\ -\,-\,-\,-\,-\,-\,-\,- \\ \vdots \\ -\,-\,-\,-\,-\,-\,-\,- \\ H_rF_{r-1}^{n-1}F_{r-2}^{n-1}\cdots F_1^{n-1} \\ H_rF_rF_{r-1}^{n-1}F_{r-2}^{n-1}\cdots F_1^{n-1} \\ \vdots \\ H_rF_r^{n-1}F_{r-1}^{n-1}\cdots F_1^{n-1} \end{bmatrix}
=\begin{bmatrix} H_1 \\ H_1F_1 \\ \vdots \\ H_1F_1^{n-1} \\ -\,-\,-\,-\,-\,-\,-\,- \\ \begin{bmatrix} H_2 \\ H_2F_2 \\ \vdots \\ H_2F_2^{n-1} \end{bmatrix} F_1^{n-1} \\ -\,-\,-\,-\,-\,-\,-\,- \\ \vdots \\ -\,-\,-\,-\,-\,-\,-\,- \\ \begin{bmatrix} H_r \\ H_rF_r \\ \vdots \\ H_rF_r^{n-1} \end{bmatrix} F_{r-1}^{n-1}F_{r-2}^{n-1}\cdots F_1^{n-1} \end{bmatrix}
=\begin{bmatrix} Q_1 \\ Q_2F_1^{n-1} \\ \vdots \\ Q_rF_{r-1}^{n-1}F_{r-2}^{n-1}\cdots F_1^{n-1} \end{bmatrix}
$$

观测性与观测度分析

4. 可观测性与可观测度的应用

3) 时变系统可观测性分析

c. 简化总的可观测性矩阵，得到可观测性矩阵(SOM)

$$Q_{SO}(r)=\begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \\ \cdots \\ Q_r \end{bmatrix}$$

根据定理(证明过程较复杂，感兴趣的可直接参考前面给出的论文)，在满足 $\mathrm{Null}\ (Q_i)\subset \mathrm{Null}\,(F_i)\quad (i=1,2,\ldots r)$ 的条件下，下式成立，

$$\mathrm{Rank}\,(Q_{TO}(r))=\mathrm{Rank}\,(Q_{SO}(r))$$

即只要计算 $Q_{SO}(r)$ 的秩，就可以判断状态量是否完全可观测。

观测性与观测度分析

4. 可观测性与可观测度的应用

4) 可观测度分析

可观测度的分析方法与前述相同，还是使用SVD进行计算。

定常系统中用来SVD分解的矩阵是 Q，时变系统中用来分解的矩阵是 Q_{SO}。

需要注意的是，在下式中，Y不是一个时刻的观测值，而是多个时刻的，以确保它和 u_i^T 的维度保持一致

$$X = \sum_{i=1}^{n} \frac{u_i^T Y v_i}{\sigma_i}$$

需要注意的是，在定位任务中，由于各状态量量纲不同（有的是速度误差，有的是姿态误差），因此需要进行归一化，修改可观测度的定义为 $\beta = \frac{\sigma_i}{\sigma_0}$，

其中 σ_i 为每个状态量对应的奇异值，而 σ_0 表示的是观测量（比如位置）对应的奇异值。当有多个观测量时，可选奇异值最大的那个。

5. 可观测性与可观测度分析总结

1) SVD分解计算可观测度的方法在理论上并不完全严谨，因此可用它做大致的判断，不必纠结细微的比较。

2) 可观测度分析还有很多其他的方式，包括kalman方差阵进行奇异值分解等。每种方式都有其缺点，SVD是目前使用最广泛的方式。

3) 时变系统在单个时刻不可观测，而多个时刻联合起来就可以观测的现象在定位任务中非常普遍，具有明确的物理意义，跟多种现象相关，比如初始对准时要绕八字、停车时精度会逐渐变差等。这类现象更多的解释会放在下一章介绍。

作业

作业1：任选一种滤波模型和方法，在kitti数据集上，实现基于点云地图的融合定位。

作业2：

1）推导组合导航(gps+imu)的滤波模型(相比于基于地图定位，只有观测方程发生了变化)，对静止、匀速、转向、加减速等不同运动状态下各状态量的可观测性和可观测度进行分析。

2) 使用第三章所述数据仿真软件，产生对应运动状态的数据，进行kalman滤波。

3) 统计kalman滤波中各状态量的收敛速度和收敛精度，并与可观测度分析的结果汇总比较。

注意：本章可观测性与可观测度分析的内容至关重要，很多现实中的现象都要靠它解释，而在解释之前，如果不能对它们和kalman估计的收敛情况之间的联系有直观认识，下一章介绍这些现象时会很难理解，因此建议大家及时完成文章作业。

感谢聆听！

Thanks for Listening