

贝叶斯滤波与卡尔曼滤波整理

在状态估计问题中，一般对系统建立状态空间模型，具体为一个状态方程（也称运动方程）和一个量测方程（也称观测方程）。

$$\begin{cases} \mathbf{x}_t = g(\mathbf{x}_{t-1}, \mathbf{u}_t) + \mathbf{w}_t \\ \mathbf{z}_t = h(\mathbf{x}_t) + \mathbf{v}_t \end{cases}$$

式中的 \mathbf{x}_t 称状态量，即表示系统在不同时刻的状态的 n 维向量； \mathbf{u}_t 称系统输入/控制量，即对系统状态进行操控的 m 维向量； \mathbf{w}_t 称系统噪声； \mathbf{z}_t 称观测量，即系统通过传感器对当前信息进行采集得到的结果； \mathbf{v}_t 称量测噪声。

由于状态量无法直接在实际系统中得到，因此需要通过系统之前的状态、输入量和观测量来推算出系统的真实状态，这个推算的过程即为状态估计。整个系统的状态估计问题可描述为，通过带有噪声的 \mathbf{u} 、 \mathbf{z} 来对载体位姿 \mathbf{x} 和地图中的路标点 \mathbf{y} 进行估计。

1. 贝叶斯滤波

若以概率的视角来看状态的演变过程，可将其视为一种条件概率分布（假定系统具有马尔可夫性，即系统当前时刻的状态仅与前一时刻的状态及当前控制量有关）。

状态转移概率（state transition probability）为

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t)$$

测量概率（measurement probability）为

$$p(\mathbf{z}_t | \mathbf{x}_t)$$

这样就构成了一个隐马尔可夫模型（Hidden Markov Model, HMM），也是一种动态贝叶斯网络（Dynamic Bayes Network, DBN）。

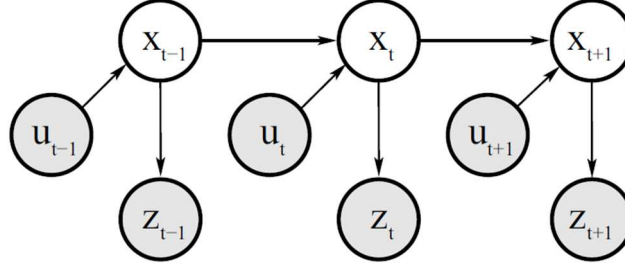


图 1 状态的演变过程

由于系统的状态无法直接测量，因此要从一个综合了量测值 \mathbf{z}_t 的后验概率（又称置信度，即 $\text{bel}(\mathbf{x}_t)$ ）中获得状态量最有可能的值。

$$\text{bel}(\mathbf{x}_t) = p(\mathbf{x}_t | \mathbf{z}_{1:t}, \mathbf{u}_{1:t})$$

在综合量测值 \mathbf{z}_t 之前的后验概率（即基于之前状态的后验概率）为

$$\overline{\text{bel}}(\mathbf{x}_t) = p(\mathbf{x}_t | \mathbf{z}_{1:t-1}, \mathbf{u}_{1:t})$$

由之前的状态计算 $\overline{\text{bel}}(\mathbf{x}_t)$ 的过程称为预测，由 $\overline{\text{bel}}(\mathbf{x}_t)$ 计算 $\text{bel}(\mathbf{x}_t)$ 的过程则称为修正。

由贝叶斯公式

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

进而有

$$\begin{aligned} p(x|y, z) &= \frac{p(y|x, z)p(x, z)}{p(y, z)} \\ &= \frac{p(y|x, z) \cdot p(x|z) \cdot p(z)}{p(y|z) \cdot p(z)} \end{aligned}$$

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

以此类推，有

$$p(x|y, z, w) = \frac{p(y|x, z, w) \cdot p(x|z, w)}{p(y|z, w)}$$

因此，状态的置信度可化为

$$\begin{aligned} \text{bel}(x_t) &= p(x_t | z_{1:t}, u_{1:t}) \\ &= p(x_t | z_t, z_{1:t-1}, u_{1:t}) \\ &= \frac{p(z_t | x_t, z_{1:t-1}, u_{1:t}) \cdot p(x_t | z_{1:t-1}, u_{1:t})}{p(z_t | z_{1:t-1}, u_{1:t})} \end{aligned}$$

由于马尔可夫性，其中 $p(z_t | z_{1:t-1}, u_{1:t})$ 无意义，可归一化为 $p^{-1}(z_t | z_{1:t-1}, u_{1:t}) = \eta$ 。那么可将 $\text{bel}(x_t)$ 化简为

$$\begin{aligned} \text{bel}(x_t) &= \eta \cdot p(z_t | x_t, z_{1:t-1}, u_{1:t}) \cdot p(x_t | z_{1:t-1}, u_{1:t}) \\ &= \eta \cdot p(z_t | x_t) \cdot p(x_t | z_{1:t-1}, u_{1:t}) \\ &= \eta \cdot p(z_t | x_t) \cdot \overline{\text{bel}}(x_t) \end{aligned}$$

就得到了修正过程。

由全概率定理

$$p(x) = \int p(x|y)p(y)dy$$

那么状态预测可化为

$$\begin{aligned} \overline{\text{bel}}(x_t) &= p(x_t | z_{1:t-1}, u_{1:t}) \\ &= \int p(x_t | x_{t-1}, z_{1:t-1}, u_{1:t}) p(x_{t-1} | z_{1:t-1}, u_{1:t}) dx_{t-1} \\ &= \int p(x_t | x_{t-1}, u_t) p(x_{t-1} | z_{1:t-1}, u_{1:t}) dx_{t-1} \end{aligned}$$

易知，

$$\overline{\text{bel}}(x_t) = \int p(x_t | x_{t-1}, u_t) \text{bel}(x_{t-1}) dx_{t-1}$$

就得到了预测过程。

综上，**贝叶斯滤波**（Bayes Filter）算法是由预测过程和修正过程组成，并不断进行迭代的一种递归的状态估计算法，如图 2 所示。

```

1:   Algorithm Bayes_filter( $\text{bel}(x_{t-1}), u_t, z_t$ ):
2:     for all  $x_t$  do
3:        $\overline{\text{bel}}(x_t) = \int p(x_t | u_t, x_{t-1}) \text{bel}(x_{t-1}) dx_{t-1}$ 
4:        $\text{bel}(x_t) = \eta p(z_t | x_t) \overline{\text{bel}}(x_t)$ 
5:     endfor
6:     return  $\text{bel}(x_t)$ 

```

图 2 贝叶斯滤波迭代过程

在迭代计算开始之前，需要给出 $t = 0$ 时的置信度 $\text{bel}(x_0)$ 作为系统初值。

值得注意的是，贝叶斯滤波器只是状态估计的一种**抽象**出的框架，仅能直接实现非常简单的估计问题。在此基础上，使用不同的假设和参数可以得到不同的滤波器，来解决各类实际问题。

2. 线性系统卡尔曼滤波

卡尔曼滤波（Kalman Filter）是贝叶斯滤波最为常用的一种具体实现。卡尔曼滤波假定系统为线性离散系统，系统噪声和量测噪声为互不相关的高斯白噪声，初始置信度服从高斯分布，以最小均方误差为最优估计准则。

则系统状态方程为

$$\mathbf{x}_t = \mathbf{A}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t + \mathbf{w}_t$$

其中系统噪声 \mathbf{w}_t 服从均值为 $\mathbf{0}$ ，方差 \mathbf{R}_t 的高斯分布，即

$$\mathbf{w}_t \sim N(\mathbf{0}, \mathbf{R}_t)$$

系统量测方程为

$$\mathbf{z}_t = \mathbf{C}_t \mathbf{x}_t + \mathbf{v}_t$$

其中量测噪声 \mathbf{v}_t 服从均值为 $\mathbf{0}$ ，方差为 \mathbf{Q}_t 的高斯分布，即

$$\mathbf{v}_t \sim N(\mathbf{0}, \mathbf{Q}_t)$$

而 \mathbf{x}_t 的初始置信度 $\text{bel}(\mathbf{x}_0)$ 服从均值为 $\boldsymbol{\mu}_0$ ，方差为 $\boldsymbol{\Sigma}_0$ 的高斯分布，即

$$\mathbf{x}_0 \sim N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$$

由高斯分布函数在经过线性变换后仍为高斯函数的性质，确保了迭代过程中，任何时刻 t 下的状态量 \mathbf{x}_t 均服从高斯分布，进而可以用均值、方差来描述随机变量的分布。卡尔曼滤波的预测和修正过程视为分别求其的均值 $\boldsymbol{\mu}$ 和方差 $\boldsymbol{\Sigma}$ 的过程。

在预测过程中，由于状态量的初值为高斯分布，那么可知状态转移概率 $p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t)$ 和 $t-1$ 时刻状态量的置信度 $\overline{\text{bel}}(\mathbf{x}_{t-1})$ 均服从高斯分布。

$$(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t) \sim N(\mathbf{A}_t \mathbf{x}_{t-1} + \mathbf{B}_t \mathbf{u}_t, \mathbf{R}_t)$$

$$\overline{\text{bel}}(\mathbf{x}_{t-1}) \sim N(\boldsymbol{\mu}_{t-1}, \boldsymbol{\Sigma}_{t-1})$$

那么预测过程的置信度为

$$\begin{aligned} \overline{\text{bel}}(\mathbf{x}_t) &= \int p(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{u}_t) \overline{\text{bel}}(\mathbf{x}_{t-1}) d\mathbf{x}_{t-1} \\ &= \int \det(2\pi\mathbf{R}_t)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_t - \mathbf{A}_t \mathbf{x}_{t-1} - \mathbf{B}_t \mathbf{u}_t)^T \mathbf{R}_t^{-1} (\mathbf{x}_t - \mathbf{A}_t \mathbf{x}_{t-1} - \mathbf{B}_t \mathbf{u}_t)\right\} \\ &\quad \det(2\pi\boldsymbol{\Sigma}_{t-1})^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_{t-1} - \boldsymbol{\mu}_{t-1})^T \boldsymbol{\Sigma}_{t-1}^{-1} (\mathbf{x}_{t-1} - \boldsymbol{\mu}_{t-1})\right\} d\mathbf{x}_{t-1} \end{aligned}$$

这实际上是两个高斯分布函数的卷积。由两个高斯函数的卷积为一新的高斯函数这一性质，可以得出预测过程的置信度的均值及方差为

$$\bar{\boldsymbol{\mu}}_t = \mathbf{A}_t \boldsymbol{\mu}_{t-1} + \mathbf{B}_t \mathbf{u}_t$$

$$\bar{\boldsymbol{\Sigma}}_t = \mathbf{A}_t \boldsymbol{\Sigma}_{t-1} \mathbf{A}_t^T + \mathbf{R}_t$$

在修正过程中，同理，测量概率及 t 时刻修正值的置信度 $\text{bel}(\mathbf{x}_t)$ 均服从高斯分布。

$$(\mathbf{z}_t | \mathbf{x}_t) \sim N(\mathbf{C}_t \mathbf{x}_t, \mathbf{Q}_t)$$

$$\text{bel}(\mathbf{x}_t) \sim N(\bar{\boldsymbol{\mu}}_t, \bar{\boldsymbol{\Sigma}}_t)$$

那么修正过程的置信度为

$$\begin{aligned} \text{bel}(\mathbf{x}_t) &= \eta \cdot p(\mathbf{z}_t | \mathbf{x}_t) \cdot \overline{\text{bel}}(\mathbf{x}_t) \\ &= \eta \cdot \det(2\pi\mathbf{Q}_t)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{z}_t - \mathbf{C}_t \mathbf{x}_t)^T \mathbf{Q}_t^{-1} (\mathbf{z}_t - \mathbf{C}_t \mathbf{x}_t)\right\} \\ &\quad \det(2\pi\bar{\boldsymbol{\Sigma}}_t)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_t - \bar{\boldsymbol{\mu}}_t)^T \bar{\boldsymbol{\Sigma}}_t^{-1} (\mathbf{x}_t - \bar{\boldsymbol{\mu}}_t)\right\} \end{aligned}$$

这实际上是两个高斯分布函数的乘积。由两个高斯函数的乘积为一新的高斯函数这一性质，可以得出修正过程的置信度的均值及方差为：

$$\begin{aligned}\mu_t &= \bar{\mu}_t + K_t(z_t - C_t \bar{\mu}_t) \\ \Sigma_t &= (I - K_t C_t) \bar{\Sigma}_t\end{aligned}$$

其中 K_t 为卡尔曼增益：

$$K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$$

综上，可以得出卡尔曼滤波算法**迭代流程**，如图 3 所示，有五个经典公式，可分为**预测过程**（2、3 行）和**修正过程**（4、5、6 行）两部分。

```

1:   Algorithm Kalman_filter( $\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$ ):
2:        $\bar{\mu}_t = A_t \mu_{t-1} + B_t u_t$ 
3:        $\bar{\Sigma}_t = A_t \Sigma_{t-1} A_t^T + R_t$ 
4:        $K_t = \bar{\Sigma}_t C_t^T (C_t \bar{\Sigma}_t C_t^T + Q_t)^{-1}$ 
5:        $\mu_t = \bar{\mu}_t + K_t(z_t - C_t \bar{\mu}_t)$ 
6:        $\Sigma_t = (I - K_t C_t) \bar{\Sigma}_t$ 
7:       return  $\mu_t, \Sigma_t$ 

```

图 3 卡尔曼滤波迭代过程

卡尔曼滤波算法五个经典公式各自的物理意义分别为：

- **第 2 行**，计算 t 时刻预测值的均值 $\bar{\mu}_t$ 。利用系统状态方程对系统状态量进行预测。
- **第 3 行**，计算 t 时刻预测值的方差 $\bar{\Sigma}_t$ 。利用系统状态方程，结合系统噪声方差，对状态量的均方误差进行预测，其大小代表了预测的精度。
- **第 4 行**，计算 t 时刻卡尔曼增益。 K_t 是使滤波估计方差达到最小的增益矩阵，同时也决定了滤波估计结果对预测值和量测值的利用权重。从公式可以看出， K_t 由量测噪声方差 Q_t 和预测置信度方差 $\bar{\Sigma}_t$ 决定。若预测精度较高，即 $\bar{\Sigma}_t$ 较小；而量测精度较低，即 Q_t 较大；那么计算出的 K_t 就会相对较小，估计结果对量测值 z_t 的权重就要小于对预测值 $\bar{\mu}_t$ 的权重。反之，则对量测值 z_t 的权重就要大于对预测值 $\bar{\mu}_t$ 的权重。考虑到预测值方差的计算过程，也可以说，卡尔曼增益 K_t 受上一时刻滤波估计方差 Σ_{t-1} 的影响，即可以根据上一时刻的信息精度，自动调整对当前信息的利用程度，具备一定的自适应性。
- **第 5 行**，计算 t 时刻修正值的均值 μ_t 。即利用量测值对预测值进行修正，以得到当前 t 时刻的状态估计值。上一步计算出的卡尔曼增益 K_t 决定了量测值和预测值之间的权重关系，即， K_t 增大，则量测值 z_t 权重增大；反之，则预测值 $\bar{\mu}_t$ 的权重增大。
- **第 6 行**，计算 t 时刻修正值的方差 Σ_t ，一定程度上反映了滤波的估计精度。

卡尔曼滤波的**直观解释**如图 4 所示。在系统具有马尔可夫性的前提下，以 $t-1$ 时刻的估计值 μ_{t-1} 为准，根据系统状态方程对系统 t 时刻的状态量进行预测，得到其预测值 $\bar{\mu}_t$ ；随后根据此时得到的观测值 z_t ，对预测值进行修正，得到修正后的值，也即是 t 时刻的估计值 μ_t 。这其中，在每次迭代过程中根据协方差计算出的卡尔曼增益 K_t 起到了一个加权求和的权值的作用，用以决定对预测值和观测值之间的信赖程度，以求得到最优的状态估计。

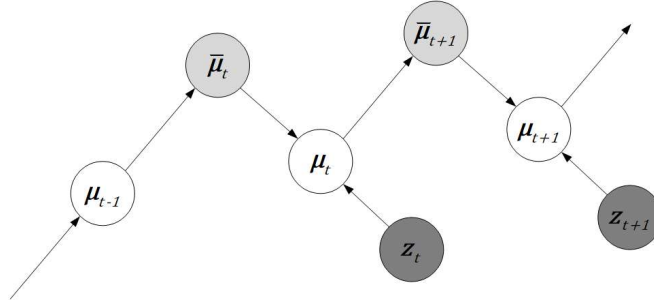


图4 卡尔曼滤波的直观示意图

3. 非线性系统卡尔曼滤波

对于具有马尔可夫性、噪声假定为高斯白噪声的**线性系统**而言，一般使用卡尔曼滤波的方式对系统状态变量进行估计；而对于**非线性系统**，高斯分布函数在经过非线性变换之后，其变换结果不再是高斯函数，因而无法使用简单的均值和方差来描述其具体的分布。很不幸，实际问题中的绝大多数系统都是非线性的。针对于这种情况，可以使用扩展卡尔曼滤波（Extended Kalman Filter, EKF）、无迹卡尔曼滤波（Unscented Kalman Filter, UKF）等方式进行状态估计。

3.1. 扩展卡尔曼滤波

扩展卡尔曼滤波的主要思想是线性化，通过在均值处选取与非线性函数相切的线性函数的方式来近似。如图5所示，这样高斯分布的状态量在经过近似的线性变换之后仍能够保持高斯分布。

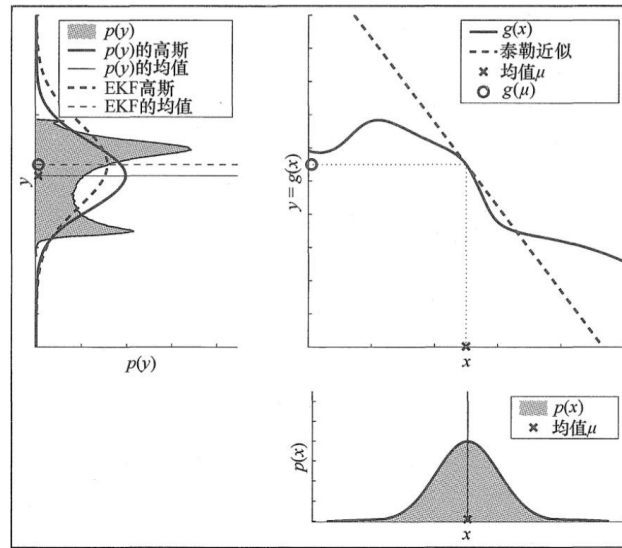


图5 扩展卡尔曼滤波基本思想

在线性化近似时，使用泰勒展开的方式寻求非线性函数的线性化。在泰勒展开的过程中，近似的精度会随着所取泰勒展开项阶数的增加而提升，但是其计算量同样会随之提升。一般情况下，受限于计算量，大多使用一阶泰勒展开来进行线性化。

将系统状态方程在 μ_{t-1} 处（状态量在该处的可能性最大）进行泰勒线性展开，可得

$$g(x_{t-1}, u_t) \approx g(\mu_{t-1}, u_t) + g'(\mu_{t-1}, u_t)(x_{t-1} - \mu_{t-1})$$

其中

$$g'(\mu_{t-1}, u_t) = \frac{\partial g(\mu_{t-1}, u_t)}{\partial x_{t-1}} \triangleq G_t$$

即， G_t 为系统状态方程的 Jacobian 矩阵。那么状态转移概率的高斯分布为

$$(x_t|x_{t-1}, u_t) \sim N(g(\mu_{t-1}, u_t) + G_t(x_{t-1} - \mu_{t-1}), R_t)$$

类似地，可以计算出预测过程中置信度的均值及方差为

$$\begin{aligned}\bar{\mu}_t &= g(\mu_{t-1}, u_t) \\ \bar{\Sigma}_t &= G_t \Sigma_{t-1} G_t^T + R_t\end{aligned}$$

而对于量测方程，可在 $\bar{\mu}_t$ 处（此时机器人认为状态量在该处的可能性最大）进行泰勒线性展开，得

$$h(x_t) \approx h(\bar{\mu}_t) + h'(\bar{\mu}_t)(x_t - \bar{\mu}_t)$$

其中

$$h'(\bar{\mu}_t) = \frac{\partial h(x_t)}{\partial x_t} \triangleq H_t$$

即， H_t 为系统量测方程的 Jacobian 矩阵。那么测量概率的高斯分布为

$$(z_t|x_t) \sim N(h(\bar{\mu}_t) + H_t(x_t - \bar{\mu}_t), Q_t)$$

之后可以计算出修正过程中置信度的均值及方差为

$$\begin{aligned}\mu_t &= \bar{\mu}_t + K_t(z_t - h(\bar{\mu}_t)) \\ \Sigma_t &= (I - K_t H_t) \bar{\Sigma}_t\end{aligned}$$

其中 K_t 为卡尔曼增益：

$$K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1}$$

综上，可以得出扩展卡尔曼滤波算法迭代流程，如图 6 所示。

```

1:   Algorithm Extended_Kalman_filter( $\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$ ):
2:      $\bar{\mu}_t = g(u_t, \mu_{t-1})$ 
3:      $\bar{\Sigma}_t = G_t \Sigma_{t-1} G_t^T + R_t$ 
4:      $K_t = \bar{\Sigma}_t H_t^T (H_t \bar{\Sigma}_t H_t^T + Q_t)^{-1}$ 
5:      $\mu_t = \bar{\mu}_t + K_t(z_t - h(\bar{\mu}_t))$ 
6:      $\Sigma_t = (I - K_t H_t) \bar{\Sigma}_t$ 
7:     return  $\mu_t, \Sigma_t$ 

```

图 6 扩展卡尔曼滤波迭代过程

对于扩展卡尔曼滤波的线性化过程而言，对于方差较大的高斯分布，泰勒展开的近似结果难以保证变换后的均值的精确程度；而对于有着较强非线性的系统，函数中无法忽略的高阶小量同样会带来较大的误差。因此，在实际应用中，状态量较小的方差、以及系统较弱的非线性程度，能够一定程度上保证算法的精度。

3.2. 无迹卡尔曼滤波

无迹卡尔曼滤波的主要思想为通过**无迹变换**（Unscented Transform, UT），来拟合高斯分布经过非线性函数后的均值和方差，以求在滤波过程中更精确的近似。

在**无迹变换**中，根据状态量 x 的均值 μ 和方差 Σ ，通过对称采样，得到加权 σ 点集 $\{x^{[i]}\}$ ，再经过非线性函数 g 的变换之后，计算变换后的均值 μ' 和方差 Σ' 。对于 n 维状态量，一般会采集 $2n + 1$ 个 σ 点，其中一个 $x^{[0]}$ 位于均值处，而其余 $2n$ 个点以均值为中心，沿各维的主轴方向对称分布于其协方差处。

$2n + 1$ 个 σ 点的选取方式为

$$\begin{aligned}\mathbf{x}^{[0]} &= \boldsymbol{\mu} \\ \mathbf{x}^{[i]} &= \boldsymbol{\mu} + \left(\sqrt{(n + \lambda)\boldsymbol{\Sigma}} \right)_i, \quad i = 1, \dots, n \\ \mathbf{x}^{[i]} &= \boldsymbol{\mu} - \left(\sqrt{(n + \lambda)\boldsymbol{\Sigma}} \right)_{i-n}, \quad i = n + 1, \dots, 2n\end{aligned}$$

每个 σ 点拥有两个权值, $w_m^{[i]}$ 用于计算均值, $w_c^{[i]}$ 用于计算方差

$$\begin{aligned}w_m^{[0]} &= \frac{\lambda}{n + \lambda} \\ w_c^{[0]} &= \frac{\lambda}{n + \lambda} + (1 - \alpha^2 + \beta) \\ w_m^{[i]} &= w_c^{[i]} = \frac{1}{2(n + \lambda)}, \quad i = 1, \dots, 2n\end{aligned}$$

其中 $\lambda = \alpha^2(n + \kappa) - n$, 而 α 、 β 、 κ 均为需要选取的参数。 α 、 κ 的取值能够决定 σ 点集的分散程度。 α 越大, 则 λ 越大, 那么 $\mathbf{x}^{[0]}$ 所占权重就越大, 而 $\mathbf{x}^{[i]}$ 在权重越小的同时距离 $\boldsymbol{\mu}$ 也越远, 一般可以取 $10^{-4} \leq \alpha \leq 1$ 。而 κ 可以取 $3 - n$, 也可以直接取 $\kappa = 0$ 。 β 的取值则与状态量 \mathbf{x} 的概率分布形式有关, 在高斯分布下, $\beta = 2$ 为最优值。

确定 σ 点的分布后, 经过非线性函数 g , 得到映射 σ 点集 $\{\mathbf{y}^{[i]}\}$

$$\mathbf{y}^{[i]} = g(\mathbf{x}^{[i]})$$

之后就可以计算变换后的均值 $\boldsymbol{\mu}'$ 和方差 $\boldsymbol{\Sigma}'$

$$\begin{aligned}\boldsymbol{\mu}' &= \sum_{i=0}^{2n} w_m^{[i]} \mathbf{y}^{[i]} \\ \boldsymbol{\Sigma}' &= \sum_{i=0}^{2n} w_c^{[i]} (\mathbf{y}^{[i]} - \boldsymbol{\mu}') (\mathbf{y}^{[i]} - \boldsymbol{\mu}')^T\end{aligned}$$

通过上述的无迹变换, 可拟合高斯分布经非线性函数变换后的均值和方差。无迹卡尔曼滤波的迭代流程如图 7 所示。其中 2-5 行为预测过程, 6-11 行用来计算相关的参数, 12-13 行为修正过程。具体的流程为:

- 第 2 行, 以 $t - 1$ 时刻估计的均值 $\boldsymbol{\mu}_{t-1}$ 和方差 $\boldsymbol{\Sigma}_{t-1}$ 为参考, 对状态量进行采样, 以确定 σ 点 \mathbf{x}_{t-1} , 其中 $\gamma = \sqrt{n + \lambda}$ 。
- 第 3 行, σ 点经过状态方程 $g(\mathbf{x}_{t-1}, \mathbf{u}_t)$, 变换为 $\bar{\mathbf{x}}_t^*$ 。
- 第 4 行, 加权计算 t 时刻预测值的均值 $\bar{\boldsymbol{\mu}}_t$ 。
- 第 5 行, 加权计算变换后的方差, 并加入系统噪声方差矩阵 \mathbf{R}_t , 结果作为 t 时刻预测值的方差 $\bar{\boldsymbol{\Sigma}}_t$ 。
- 第 6 行, 以 t 时刻预测的均值 $\bar{\boldsymbol{\mu}}_t$ 和方差 $\bar{\boldsymbol{\Sigma}}_t$ 为参考, 对此时的状态量进行采样, 得到新的 σ 点 $\bar{\mathbf{x}}_t$ 。
- 第 7 行, σ 点经过量测方程 $h(\mathbf{x}_t)$, 变换为 $\bar{\mathbf{z}}_t$ 。
- 第 8 行, 加权计算变换后的均值 $\hat{\mathbf{z}}_t$, 作为预测观测值, 加入到后面的计算中。
- 第 9 行, 加权计算变换后的方差, 并加入当前量测噪声方差矩阵 \mathbf{Q}_t , 结果记为 \mathbf{S}_t 。作为预测观测值的不确定性加入到后面的计算中。
- 第 10 行, 加权计算变换前与变换后 σ 点的协方差, 结果记为 $\bar{\boldsymbol{\Sigma}}_t^{x,z}$, 加入后面的计算中。
- 第 11 行, 计算 t 时刻卡尔曼增益 \mathbf{K}_t 。
- 第 12 行, 计算 t 时刻状态量修正值的均值, 即状态估计的值。
- 第 13 行, 计算 t 时刻状态量修正值的方差, 即状态估计的方差。

```

1:   Algorithm Unscented_Kalman_filter( $\mu_{t-1}, \Sigma_{t-1}, u_t, z_t$ ):
2:        $\mathcal{X}_{t-1} = (\mu_{t-1} \quad \mu_{t-1} + \gamma\sqrt{\Sigma_{t-1}} \quad \mu_{t-1} - \gamma\sqrt{\Sigma_{t-1}})$ 
3:        $\bar{\mathcal{X}}_t^* = g(u_t, \mathcal{X}_{t-1})$ 
4:        $\bar{\mu}_t = \sum_{i=0}^{2n} w_m^{[i]} \bar{\mathcal{X}}_t^{*[i]}$ 
5:        $\bar{\Sigma}_t = \sum_{i=0}^{2n} w_c^{[i]} (\bar{\mathcal{X}}_t^{*[i]} - \bar{\mu}_t)(\bar{\mathcal{X}}_t^{*[i]} - \bar{\mu}_t)^T + R_t$ 
6:        $\bar{\mathcal{X}}_t = (\bar{\mu}_t \quad \bar{\mu}_t + \gamma\sqrt{\bar{\Sigma}_t} \quad \bar{\mu}_t - \gamma\sqrt{\bar{\Sigma}_t})$ 
7:        $\bar{\mathcal{Z}}_t = h(\bar{\mathcal{X}}_t)$ 
8:        $\hat{z}_t = \sum_{i=0}^{2n} w_m^{[i]} \bar{\mathcal{Z}}_t^{[i]}$ 
9:        $S_t = \sum_{i=0}^{2n} w_c^{[i]} (\bar{\mathcal{Z}}_t^{[i]} - \hat{z}_t)(\bar{\mathcal{Z}}_t^{[i]} - \hat{z}_t)^T + Q_t$ 
10:       $\Sigma_t^{x,z} = \sum_{i=0}^{2n} w_c^{[i]} (\bar{\mathcal{X}}_t^{[i]} - \bar{\mu}_t)(\bar{\mathcal{Z}}_t^{[i]} - \hat{z}_t)^T$ 
11:       $K_t = \Sigma_t^{x,z} S_t^{-1}$ 
12:       $\mu_t = \bar{\mu}_t + K_t(z_t - \hat{z}_t)$ 
13:       $\Sigma_t = \bar{\Sigma}_t - K_t S_t K_t^T$ 
14:      return  $\mu_t, \Sigma_t$ 

```

图 7 无迹卡尔曼滤波迭代流程

在强非线性的条件下，无迹卡尔曼滤波在避免了 Jacobian 矩阵的复杂计算的同时，其精度也要好于一阶泰勒展开的扩展卡尔曼滤波。因此，在弱非线性的条件下，扩展卡尔曼滤波更加适用；而在强非线性的条件下，无迹卡尔曼滤波则更加适用。

参考

- [1] 高翔, 张涛, 刘毅, 等. 视觉 SLAM 十四讲[M]. 北京: 电子工业出版社, 2017.
- [2] Thrun S, Burgard W, Fox D. 概率机器人[M]. 曹红玉, 谭志, 史晓霞, 等译. 北京: 机械工业出版社, 2017.
- [3] 秦永元, 张洪钺, 汪淑华. 卡尔曼滤波与组合导航原理[M]. 第 3 版. 西安:西北工业大学出版社, 2015.
- [4] 刘胜, 张红梅. 最优估计理论[M]. 北京:科学出版社, 2011.