概率机器人笔记

第一章 数学基础

一、概率论

【1.1 条件概率公式】

1. 定义

$$p(A|B) = \frac{p(AB)}{P(B)} \rightarrow p(AB) = p(A|B) \cdot p(B) \tag{1}$$

2. 如果将两件事看作一件事

$$\begin{aligned} p(B|A,C) &= \frac{p(ABC)}{p(AC)} \\ &= \frac{p(A|B,C) \cdot p(BC)}{p(AC)} \\ &= \frac{p(A|B,C) \cdot p(B|C) \cdot p(C)}{p(A|C) \cdot p(C)} \\ &= \frac{p(A|B,C) \cdot p(B|C)}{p(A|C)} \end{aligned} \tag{2}$$

$$p(A|B,C) = \frac{p(B|A,C) \cdot p(A|C)}{p(B|C)} \tag{3}$$

3. **条件独立：**$p(x,y|z) = p(x|z)p(y|z)$，表达以z为条件的两个独立随机变量的联合概率，并且，条件独立定律一般不能和独立时间公式$p(x,y) = p(x)p(y)$互推。

【1.2 贝叶斯】

1. 贝叶斯（Bayes）的先验概率：如果x是希望从y中推测出来的数值，则$p(x)$称为**先验概率**，先验概率不包含任何关于从y到x的推测，仅仅是人们对x的经验认识，y称为**数据**，在机器人中就是传感器测量值。$p(x)$总结了在综合数据y之前已经有的关于x的信息。
2. 贝叶斯的后验概率：从传感器数据y推断x的值得到的关于x的概率分布$p(x|y)$，称为**在X上的后验概率分布**。由条件概率公式（贝叶斯定理），后验概率分布$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$，由于分母$p(y)$不依赖任何$x$，因此对于任何$x$，$p(y)$总是常数，称为**归一化常数**。所以，贝叶斯定理的简洁表达式为：$p(x|y) = \eta p(y|x)p(x)$。

【1.3方差、协方差、协方差矩阵】

1. 方差：一个随机变量X分布的离散度

$$\begin{aligned}\sigma^2 &= D(X) \qquad (4)\\ &= E\{[X-E(X)]^2\}\\ &= \overline{(X-\overline{X})^2)}\\ &= E(X^2)-[E(X)]^2 \qquad (5)\\ &= \overline{X^2}-\overline{X}^2 \qquad (6)\end{aligned}$$

2. 协方差：两个随机变量X和Y之间的线性相关性（协方差为零也有可能存在非线性相关性）

$$\operatorname{Cov}(X,Y)=E\{[X-E(X)][Y-E(Y)]\} \tag{7}$$

（线性）相关系数：

$$\rho_{xy}=\frac{\operatorname{Cov}(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} \tag{8}$$

3. 协方差矩阵：若干个随机变量$\boldsymbol{x}=\{x_1,x_2,\cdots,x_n\}$之间的线性相关性

$$\boldsymbol{C}=\begin{pmatrix}\operatorname{Cov}(x_1,x_1),\operatorname{Cov}(x_1,x_2),\cdots,\operatorname{Cov}(x_1,x_n)\\ \operatorname{Cov}(x_2,x_1),\operatorname{Cov}(x_2,x_2),\cdots,\operatorname{Cov}(x_2,x_n)\\ \vdots,\vdots,\ddots,\vdots,\\ \operatorname{Cov}(x_n,x_1),\operatorname{Cov}(x_n,x_2),\cdots,\operatorname{Cov}(x_n,x_n)\end{pmatrix} \tag{37}$$

【1.4其他】

1. 多元正态分布：$p(\boldsymbol{x})=\det(2\pi\Sigma)^{-\frac{1}{2}}\exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}\boldsymbol{x}-\boldsymbol{\mu})\}$
2. 一个概率分布$p(x)$的熵：

$$\begin{aligned}H_p(x) &= E(-\log_2 p(x)) \qquad (10)\\ H_p(x) &= -\int p(x)\log_2 p(x)\mathrm{d}x \quad \text{（连续）} \qquad (11)\\ H_p(x) &= -\sum_x p(x)\log_2 p(x) \quad \text{（离散）} \qquad (12)\end{aligned}$$

二、线性代数

1. 二次型：关于向量$\boldsymbol{x}=\{x_1,x_2,\cdots,x_n\}$中各个元素的二次函数，可以用矩阵表达。

第二章 递归状态估计

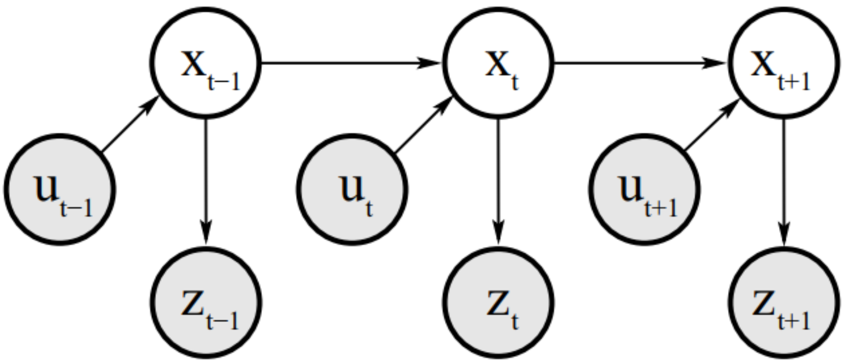
1. 状态：包含环境（动态状态：如周围人物的走动等，静态状态：建筑物的位置等）和机器人本身（机器人的位姿，速度、传感器是否正常等）的会对未来产生影响的所有方面的因素，用$\boldsymbol{x}$表示。一般包括
   - 机器人的位姿：机器人相对于全局坐标系的位置和方向

- 机器人执行机构配置（运动学状态）：各个关节的关节角度以及其它自由度
- 机器人的运动状态：速度、角速度
- 机器人的环境状态：环境中周围物体的位置和特征

2. 环境交互：包含**环境传感器测量**和**控制动作改变世界地状态**两部分。机器人连续地执行控制，同时进行测量。

3. 概率生成法则，状态和测量的演变由概率法则支配，状态转移概率和测量概率仪器描述机器人及其环境组成的动态随机系统。

   - 状态转移概率：$p(\boldsymbol{x}_t|\boldsymbol{x}_{0:t-1},\boldsymbol{z}_{0:t-1},\boldsymbol{u}_{1:t})=p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{u}_t)$，条件独立公式（条件独立的含义没理解），即遵循马尔科夫假设；
   - 测量概率：$p(\boldsymbol{z}_t|\boldsymbol{x}_t)$或者$p(\boldsymbol{z}|\boldsymbol{x})$，表明如果$\boldsymbol{x}$是完整的，则状态$\boldsymbol{x}_t$足以预测（有潜在噪声的）测量$\boldsymbol{z}_t$，过去的测量、控制、状态等其他变量信息都与之无关。



4. 置信度（信息的状态/信息状态）-置信分布

   - 置信度分布是**以可获得数据（所有过去的测量z_t和所有过去的控制u_t）为条件**的**关于状态变量的**后验概率$\mathrm{bel}(\boldsymbol{x}_t)=p(\boldsymbol{x}_t|\boldsymbol{z}_{1:t},\boldsymbol{u}_{1:t})$，注意$z$的下表是$1:t$而不是$1:t-1$。
   - 综合$\boldsymbol{z}_t$之前计算的后验概率或者置信度（有些地方叫做似然）$\overline{\mathrm{bel}}(\boldsymbol{x}_t)=p(\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})$，是通过$t$之前历史测量数据对状态转移概率（先验）的修正。

5. 贝叶斯（Bayes）滤波

   贝叶斯滤波使用上一个时刻$t-1$的状态置信度，来推测当前时刻t的状态置信度。假设有一个系统递归方程$\boldsymbol{x}_t=f(\boldsymbol{x}_{t-1})+\boldsymbol{v}(t)$，其中$\boldsymbol{v}(t)$是噪声，和一个测量方程$\boldsymbol{z}_t=g(\boldsymbol{x}_t+\boldsymbol{n}_t)$。

   首先通过系统方程可以得到当前时刻的状态转移概率，即先验概率分布（注意状态分布和噪声分布一样，再加上平移）$p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})=p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{u}_t)$，此处采用马尔科夫假设（状态完整性假设）。

   另一方面，既然已经有了$[1,t-1]$时刻的测量数据，可以使用历史测量数据对t时刻的状态概率分布进行估计$p(\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})$，根据联合分布公式可得

$$\overline{\mathrm{bel}}(\boldsymbol{x}_t)=p(\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})=\int p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})\cdot p(\boldsymbol{x}_{t-1}|\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})\mathrm{d}\boldsymbol{x}_{t-1} \quad (13)$$

   (13)式也就是综合z_t之前计算的置信度，有些地方叫做似然，其中积分符号内是状态转移概率与前一个时刻的状态置信度的乘积。前者通过系统方程得到，后者为已知。**该式子也可以理解为用$t-1$时刻的状态置信度对t时刻的状态转移概率求期望值，期望值就是在t时刻上结合了历史测量数据的状态分布预测，也就是$\overline{\mathrm{bel}}(\boldsymbol{x}_t)$。**

   然后，通过测量方程得到当前时刻系统测量结果及其分布$p(\boldsymbol{z}_t|\boldsymbol{x}_t,\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})=p(\boldsymbol{z}_t|\boldsymbol{x}_t)$。

   当前时刻系统状态分布的置信度$\mathrm{bel}(\boldsymbol{x}_t)=p(\boldsymbol{x}_t|\boldsymbol{z}_{1:t},\boldsymbol{u}_{1:t})$，可以根据条件概率公式（贝叶斯准则），即(3)式推导得出

$$
\begin{aligned}
\mathrm{bel}(\boldsymbol{x}_t) &= p(\boldsymbol{x}_t|\boldsymbol{z}_{1:t},\boldsymbol{u}_{1:t}) \\
&= p(\boldsymbol{x}_t|\boldsymbol{z}_t,\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t}) \\
&= \frac{p(\boldsymbol{z}_t|\boldsymbol{x}_t,\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})\cdot p(\boldsymbol{x}_t|\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})}{p(\boldsymbol{z}_t|\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})} \\
&= \eta p(\boldsymbol{z}_t|\boldsymbol{x}_t)\cdot\overline{\mathrm{bel}}(\boldsymbol{x}_t)
\end{aligned} \tag{14}
$$

其中$\eta = p(\boldsymbol{z}_t|\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t})^{-1}$与$\boldsymbol{x}_t$无关，为归一化常数，可以通过归一化方法来计算，或者使用积分方法计算

$$
\eta = p(\boldsymbol{z}_t|\boldsymbol{z}_{1:t-1},\boldsymbol{u}_{1:t}) = \int p(\boldsymbol{z}_t|\boldsymbol{x}_{t-1})\mathrm{d}\boldsymbol{x}_{t-1} \tag{15}
$$

通过(14)式将当前时刻的系统状态置信度变为测量概率分布与**似然**的乘积，而似然又可以根据前一个时刻的置信度与状态转移概率得出，当$t=1$时，需要初始置信度$\mathrm{bel}(\boldsymbol{x}_0)$。这就是贝叶斯滤波方法，其中(1)~(2)为预测，(4)为测量，(5)为更新。

贝叶斯滤波算法的步骤：

- (0) 初始化初始置信度$\mathrm{bel}(\boldsymbol{x}_0)$，$t=1$；
- **(1) 根据系统方程计算状态转移概率作为先验概率分布$p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{u}_t)$；**
- **(2) 根据状态转移概率$p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1},\boldsymbol{u}_t)$和前一个时刻的状态置信度$\mathrm{bel}(\boldsymbol{x}_{t-1})$，利用(13)式计算$\overline{\mathrm{bel}}(\boldsymbol{x}_t)$；**
- **(3) 根据(15)式计算t时刻的归一化常数η_t；**
- **(4) 根据系统测量方程得到测量概率分布$p(\boldsymbol{z}_t|\boldsymbol{x}_t)$；**
- **(5) 根据测量测量概率分布、似然（即综合z_t之前的置信度）以及归一化常数，利用(14)式计算置信度；**
- (6) $t=t+1$，重复（2）到（6），直到满足退出条件。

第三章 高斯滤波

高斯滤波家族是实现贝叶斯滤波的一类方式，具体的还分为不同的方法，如卡尔曼滤波、扩展卡尔曼滤波、信息滤波等。所有高斯滤波都用多元高斯分布表示置信度

$$
p(\boldsymbol{x}) = \det(2\pi\Sigma)^{-\frac{1}{2}}\exp\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}\boldsymbol{x}-\boldsymbol{\mu})\} \tag{16}
$$

1. 卡尔曼滤波（Kalman Filter, KF）
   - 卡尔曼滤波用**矩参数，即概率分布的一阶矩（期望）、二阶矩（协方差矩阵）、高阶矩**，表示置信度。
   - 高斯后验，贝叶斯滤波中，如果除马尔科夫假设之外还有如下三个特性，则后验就是高斯的：
     - 状态转移概率必须是**带有随机高斯噪声的参数**的线性函数，其系统方程可以表示为(17)式，其中ε_t是一个高斯随机变量，表示状态转移的不确定性，其维数和状态向量维数相同，期望为$\mathbf{0}$，协方差矩阵为$\boldsymbol{R}_t$。这样的系统称为**线性高斯系统**，反映了它带有附加高斯噪声的自变量呈线性。
     $$\boldsymbol{x}_t = \boldsymbol{A}_t\boldsymbol{x}_{t-1} + \boldsymbol{B}_t\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t \tag{17}$$
     - 测量概率也**与带有高斯噪声的自变量呈线性关系**，其测量方程可以表示为(18)式，其中$\boldsymbol{\delta}_t$为测量噪声，为高斯分布，期望式$\mathbf{0}$，协方差矩阵为（后都称作协方差）$\boldsymbol{Q}_t$。
     $$\boldsymbol{z}_t = \boldsymbol{C}_t\boldsymbol{x}_t + \boldsymbol{\delta}_t \tag{18}$$
     - 初始置信度必须是高斯分布，期望为$\boldsymbol{\mu}_0$，协方差为$\boldsymbol{\Sigma}_t$。

   这三个假设保证了置信度在任何时刻都是高斯的。
   - 数学原理

卡尔曼滤波也从贝叶斯滤波出发，其流程和贝叶斯滤波相同，经过了预测和测量更新两步。其中预测部分求出了$\overline{\mathrm{bel}}(\boldsymbol{x}_t)$，更新部分求出$\mathrm{bel}(\boldsymbol{x}_t)$。由于两者都呈高斯分布，因此，对于分布的计算只需求出其**期望μ和协方差$\boldsymbol{\Sigma}$**。这里说明一下多元高斯分布的含义，多元高斯分布(16)其指数是关于$\boldsymbol{x}$的二次型，即指数部分如果能表示呈关于某一个向量的二次型，则该函数就是高斯分布。形如$p(\boldsymbol{x}) = \eta \exp[-\mathscr{L}(\boldsymbol{x})]$，其中$\mathscr{L}(\boldsymbol{x})$是二次型的归一化函数就是高斯分布。高斯分布有两个性质：1. $\boldsymbol{\mu} = \arg[\frac{\partial \mathscr{L}}{\partial \boldsymbol{x}} = 0]$，2. $\boldsymbol{\Sigma} = [\frac{\partial^2 \mathscr{L}}{\partial \boldsymbol{x}^2}]^{-1}$。

- **预测：** 根据(13)式以及线性高斯假设，

$$\begin{aligned}\overline{\mathrm{bel}}(\boldsymbol{x}_t) &= \int \mathcal{N}(\boldsymbol{x}_t; \mu_{t-1\to t}, \Sigma_{t-1\to t}) \mathcal{N}(\boldsymbol{x}_{t-1}; \mu_{t-1}, \Sigma_{t-1}) \mathrm{d}\boldsymbol{x}_{t-1} && (19)\\ &= \mathcal{N}(\boldsymbol{x}_t; \bar{\boldsymbol{\mu}}_t, \bar{\boldsymbol{\Sigma}}_t) && (20)\end{aligned}$$

  其中$\mathcal{N}(x; \mu, \sigma^2)$表示高斯分布，$\mu_{t-1\to t}$表示状态转移概率的期望，$\Sigma_{t-1\to t}$表示状态转移概率的协方差，$\mu_{t-1}$表示$t-1$时刻的置信度期望，$\Sigma_{t-1}$表示$t-1$时刻置信度的协方差。(19)式是两个高斯分布乘积的积分，其结果仍然是高斯分布。**由于两个高斯分布的乘积，其各自的指数部分都是关于各自随机变量的二次型，指数函数相乘即指数相加，二次型相加仍然为各自随机变量的新二次型，新二次型记为$\mathscr{L}(\boldsymbol{x}_t)$。积分积掉$\boldsymbol{x}_{t-1}$，对$\boldsymbol{x}_t$仍然为高斯分布，对$\mathscr{L}(\boldsymbol{x}_t)$求一阶导数和二阶导数即可求出$\bar{\mu}$和$\bar{\boldsymbol{\Sigma}}_t$。**

$$\begin{aligned}\bar{\boldsymbol{\mu}} &= \boldsymbol{A}_t \boldsymbol{\mu}_{t-1} + \boldsymbol{B}_t \boldsymbol{u}_t && (21)\\ \bar{\boldsymbol{\Sigma}}_t &= (\boldsymbol{A}_t \boldsymbol{\Sigma}_{t-1}^{-1} \boldsymbol{A}_t^{\mathrm{T}} + \boldsymbol{R}_t)^{-1} && (22)\end{aligned}$$

- **更新：** 根据(14)式以及线性高斯假设，

$$\begin{aligned}\mathrm{bel}(\boldsymbol{x}_t) &= \eta \mathcal{N}(\boldsymbol{z}_t; \boldsymbol{\mu}_t^z, \boldsymbol{Q}_t) \cdot \mathcal{N}(\boldsymbol{x}_t; \bar{\boldsymbol{\mu}}_t \bar{\boldsymbol{\Sigma}}_t) && (23)\\ &= \mathcal{N}(\boldsymbol{x}_t; \boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t) && (24)\end{aligned}$$

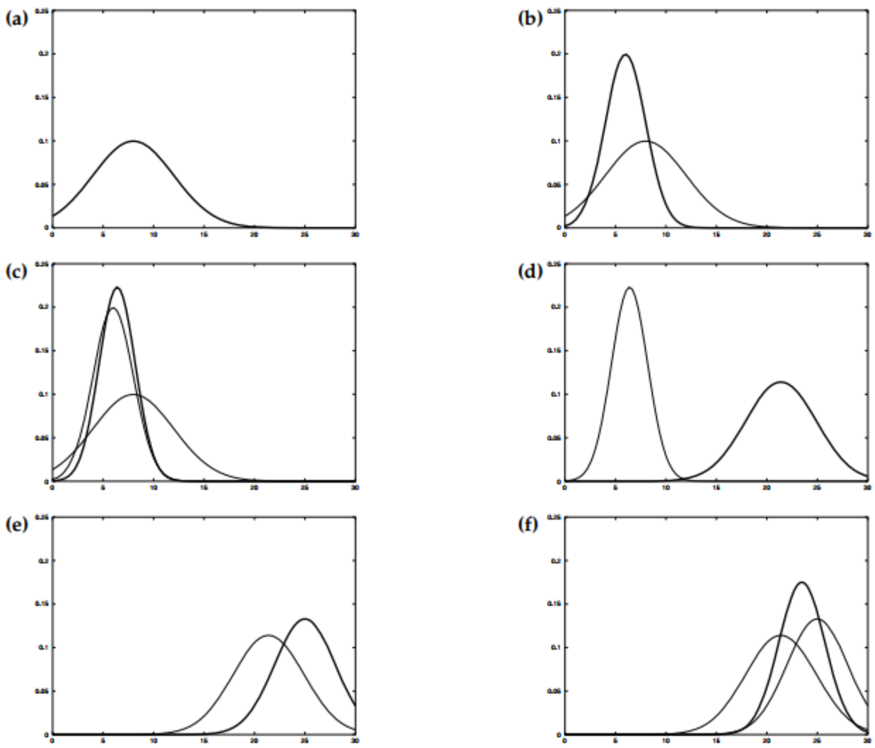
  其中，$\boldsymbol{\mu}_t^z = \boldsymbol{C}_t \boldsymbol{x}_t$为测量期望，$\boldsymbol{Q}_t$为测量协方差，$\boldsymbol{\mu}_t$为$t$时刻置信度期望，$\boldsymbol{\Sigma}_t$为$t$时刻置信度协方差。**两个高斯分布相乘同样是高斯分布。**

$$\begin{aligned}\boldsymbol{K}_t &= \boldsymbol{\Sigma}_t \boldsymbol{C}_t^{\mathrm{T}} \boldsymbol{Q}_t^{-1}\\ &= \bar{\boldsymbol{\Sigma}}_t \boldsymbol{C}_t^{\mathrm{T}} (\boldsymbol{C}_t \bar{\boldsymbol{\Sigma}}_t \boldsymbol{C}_t^{\mathrm{T}} + \boldsymbol{Q}_t)^{-1} && (25)\end{aligned}$$

$$\begin{aligned}\boldsymbol{\mu}_t &= \bar{\boldsymbol{\mu}}_t + \boldsymbol{K}_t (\boldsymbol{z}_t - \boldsymbol{C}_t \bar{\boldsymbol{\mu}}_t) && (26)\\ \boldsymbol{\Sigma}_t &= (\boldsymbol{I} - \boldsymbol{K}_t \boldsymbol{C}_t) \bar{\boldsymbol{\Sigma}}_t && (27)\end{aligned}$$

  其中，$\boldsymbol{K}_t$为卡尔曼增益。这就计算出了当前时刻的系统状态置信度，只需要按照顺序计算五个红色公式即可。

- 书上的例子



(a)先验和初始置信度；(b)测量；(c)置信度；(d)状态转移（机器人移动），先验和上一步置信度结合$\overline{\mathrm{bel}}(\boldsymbol{x}_t)$；(e)测量；(f)置信度。

2. 扩展卡尔曼滤波（Extend Kalman Filter, EKF）

卡尔曼滤波的重要假设是线性高斯系统，其中**线性**的含义是，**①观测是状态的线性函数**，并且②**下一个状态是前一个状态的线性函数。高斯线性系统的特性是，高斯随机变量的任何线性变换都将导致另一个高斯随机变量。**不幸的是，实际中状态转移和测量很少是线性的，这使得卡尔曼滤波算法不适用于除了最平凡的机器人问题以外的其他所有问题（好象很严峻的样子）。

扩展卡尔曼滤波假设状态转移概率和测量概率分别由非线性函数g和h控制。即

$$\boldsymbol{x}_t = g(\boldsymbol{u}_t, \boldsymbol{x}_{t-1}) + \boldsymbol{\varepsilon}_t \tag{28}$$

$$\boldsymbol{z}_t = h(\boldsymbol{x}_t) + \boldsymbol{\delta}_t \tag{29}$$

准确地实现对置信度地更新对非线性函数是不可能的，因此扩展卡尔曼滤波计算真实置信度的高斯近似，其期望和协方差仍然为$\boldsymbol{\mu}_t$和$\boldsymbol{\Sigma}_t$，不同的是，扩展卡尔曼滤波的目标就从计算精确的后验概率转变为有效地估计其期望和协方差。

扩展卡尔曼滤波通过一个与非线性函数g在高斯均值处相切地线性函数取近似非线性函数g。

- 通过泰勒展开线性化
  - 将函数$g(\boldsymbol{u}_t, \boldsymbol{x}_{t-1})$在$\boldsymbol{x}_{t-1} = \boldsymbol{\mu}_{t-1}$处（认为在$t-1$时刻，机器人最有可能的状态就是$\boldsymbol{\mu}_{t-1}$）做一阶泰勒展开

$$\begin{aligned} g(\boldsymbol{u}_t, \boldsymbol{x}_{t-1}) &= g(\boldsymbol{u}_t, \boldsymbol{\mu}_{t-1}) + \frac{\partial g(\boldsymbol{u}_t, \boldsymbol{x}_{t-1})}{\partial \boldsymbol{x}_{t-1}}\Big|_{\boldsymbol{x}_{t-1}=\boldsymbol{\mu}_{t-1}} + \cdots \\ &\approx g(\boldsymbol{u}_t, \boldsymbol{\mu}_{t-1}) + \boldsymbol{G}_t(\boldsymbol{x}_{t-1} - \boldsymbol{\mu}_{t-1}) \end{aligned} \tag{30}$$

    (30)式就是一个线性函数，和(17)式一样，满足了**线性要求中**，**下一个状态是前一个状态的线性函数**这一要求，$\boldsymbol{G}_t = \frac{\partial g(\boldsymbol{u}_t, \boldsymbol{x}_{t-1})}{\partial \boldsymbol{x}_{t-1}}\Big|_{\boldsymbol{x}_{t-1}=\boldsymbol{\mu}_{t-1}} = ?$，是雅可比矩阵。

  - 同理，将函数$h(\boldsymbol{x}_t)$在$\boldsymbol{x}_t = \bar{\boldsymbol{\mu}}_t$处做一阶泰勒展开

$$\begin{aligned} h(\boldsymbol{x}_t) &= h(\bar{\boldsymbol{\mu}}_t) + h'(\bar{\boldsymbol{\mu}}_t)(\boldsymbol{x}_t - \bar{\boldsymbol{\mu}}_t) + \cdots \\ &\approx h(\bar{\boldsymbol{\mu}}_t) + \boldsymbol{H}_t(\boldsymbol{x}_t - \bar{\boldsymbol{\mu}}_t) \end{aligned} \tag{31}$$

满足了**线性要求中**，**观测是状态的线性函数**这一要求，$\boldsymbol{H}_t$是雅可比矩阵。

- 扩展卡尔曼滤波算法：满足了系统的线性要求，只要初始置信度是高斯的，则就可以使用卡尔曼滤波数学原理，得出扩展卡尔曼滤波的算法过程，即

$$\bar{\boldsymbol{\mu}}_t = g(\boldsymbol{u}_t, \boldsymbol{\mu}_{t-1}) \tag{32}$$
$$\bar{\boldsymbol{\Sigma}}_t = \boldsymbol{G}_t \boldsymbol{\Sigma}_{t-1} \boldsymbol{G}_t^{\mathrm{T}} + \boldsymbol{R}_t \tag{33}$$

$$\boldsymbol{K}_t = \bar{\boldsymbol{\Sigma}}_t \boldsymbol{H}_t^{\mathrm{T}} (\boldsymbol{H}_t \bar{\boldsymbol{\Sigma}}_t \boldsymbol{H}_t^{\mathrm{T}} + \boldsymbol{Q}_t)^{-1} \tag{34}$$

$$\boldsymbol{\mu}_t = \bar{\boldsymbol{\mu}}_t + \boldsymbol{K}_t (\boldsymbol{z}_t - h(\bar{\boldsymbol{\mu}}_t)) \tag{35}$$
$$\boldsymbol{\Sigma}_t = (\boldsymbol{I} - \boldsymbol{K}_t \boldsymbol{H}_t) \bar{\boldsymbol{\Sigma}}_t \tag{36}$$

- 局限：扩展卡尔曼滤波所应用的线性近似是否有优势取决于两个因素：被近似的**局部线性化程度**和**不确定度**，较高的不确定度通常会导致结果随机变量的均值和协方差估计更不精确，而更强的g的局部非线性会产生更大的近似误差。

3. 无迹卡尔曼滤波（Unscented Kalman Filter, UKF）

   无迹卡尔曼滤波采用另一种线性化方法，它通过使用加权统计线性回归过程实现随机线性化。