卡尔曼增益推导

一、卡尔曼情景回顾

1. 运动模型与观测模型:

$$\begin{cases} x_k = f(x_{k-1}, u_k + \omega_k) \\ z_{k,j} = h(h_j, x_k) \end{cases}$$

其中 f 和 h 分别为运动模型和观测模型的非线性函数。运动模型和观测模型的噪声分别为 ω_k 和 v_k 。其中 x_k 为系统位姿， y_k 为路标点， u_k 为控制输入， $z_{k,j}$ 为观测值， h_j 为观测模型的参数。

而后端就是从概率论的角度，根据已有的观测值和运动数据，得到系统位姿 x_k 和路标点 y_k 的后验概率分布。这里我们比较关注位姿的分布，即$P(x_k|x_0, u_{1...k}, z_{1...k})$。

为了与上述的运动模型和观测模型相对应，我们可以利用贝叶斯公式，得到:

$$P(x_k|x_0, u_{1...k}, z_{1...k}) = \frac{P(x_k, x_0, u_{1...k}, z_{1...k})}{P(x_0, u_{1...k}, z_{1...k})} = \frac{P(z_k|x_{k|x_0, u_{1...k}, z_{1...k-1}})P(x_k|x_0, u_{1...k}, z_{1...k-1})}{P(x_0, u_{1...k}, z_{1...k})}$$

其中分母为已知数据，因此概率为常数。我们可以将其省略掉；分子中$P(x_k)$ 即 $P(x_k|x_0, u_{1...k}, z_{1...k-1})$, 因此得到下面的式子，也就是两个方程的概率乘积

$$P(x_k|x_0, u_{1...k}, z_{1...k}) = P(z_k|x_k)P(x_k|x_0, u_{1...k}, z_{1...k-1})$$

2. 线性化模型:

$$\begin{cases} x_k = A_k x_{k-1} + u_k + \omega_k \\ z_k = C_k x_k + v_k \end{cases}$$

为简化起见，我们将运动模型和观测模型线性化，得到上面的线性化模型。同时我们假设了马尔可夫性：当前状态的位姿仅仅与前一时刻的状态有关。

其中， A_k 和 C_k 分别为运动模型和观测模型的线性矩阵。我们假设噪声服从高斯分布： ω_k是运动模型的噪声， v_k是观测模型的噪声。

$$\omega_k \sim N(0, R), v_k \sim N(0, Q)$$

二、卡尔曼增益推导

1. 预测部分:

我们将两个方程代入这个两个高斯分布，于是两个方程被转化为两个分布：

$$\begin{cases} P(x_k|x_0, u_{1...k}, z_{1...k-1}) = N(A_k \hat{x}_{k-1} + u_k, A_k \hat{P}_{k-1} A_k^T + R) \\ P(z_k|x_k) = N(C_k x_k, Q) \end{cases}$$

请注意，观测方程中的 x_k 是已经确定的状态，而不是随机变量。我们再记先验状态：

$$\check{x}_k = A_k\hat{x}_{k-1} + u_k, \quad \check{P}_k = A_k\hat{P}_{k-1}A_k^T + R$$

我们可以将上面的两个分布代入后验公式中，得到：

$$P(x_k|x_0, u_{1\ldots k}, z_{1\ldots k}) = \eta N(C_k x_k, Q), N(\check{x}_k, \check{P}_k)$$

2. 更新部分:

接下来还是和书中一样，比较指数部分的二次型配方，无需处理高斯分布的因子部分，将指数部分展开，得到

$$(x_k - \hat{x}_k)^T\hat{P}_k^{-1}(x_k - \hat{x}_k) = (z_k - C_k x_k)^T Q^{-1}(z_k - C_k x_k) + (x_k - \check{x}_k)^T\check{P}_k^{-1}(x_k - \check{x}_k)$$

通过比较x_k的二次项和一次项系数我们可以分别得到后验分布的均值和方差。首先是二次项，进行同乘$\hat{P}_k$并移项的处理：

$$\hat{P}_k^{-1} = C_k^T Q^{-1} C_k + \check{P}_k^{-1}$$

$$\check{P}_k^{-1}\hat{P}_k = \hat{P}_k^{-1}\hat{P}_k - C_k^T Q^{-1} C_k\hat{P}_k$$

在这里，我们引入卡尔曼增益：

$$K = \hat{P}_k C_k^T Q^{-1}$$

因此我们可以将其代入前一个式子，得到后验方差：

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

接下来是一次项：

$$-2\hat{x}_k^T\hat{P}_k^{-1}x_k = -2z_k^T Q^{-1} C_k x_k - 2\check{x}_k^T\check{P}_k^{-1}x_k$$

我们可以将其变形去掉无关的系数，再代入后验方差，并同乘$\hat{P}_k$，得到后验均值：

$$\hat{P}_k^{-1}\hat{x}_k = Q^{-1}C_k z_k + \check{P}_k^{-1}\check{x}_k$$

$$\hat{x}_k = \hat{P}_k(Q^{-1}C_k z_k + \check{P}_k^{-1}\check{x}_k)$$

$$= Kz_k + (I - KC_k)\check{x}_k = \check{x}_k + K(z_k - C_k\check{x}_k)$$

至此，我们得到了后验均值和方差：

$$\begin{cases} \hat{P}_k = (I - KC_k)\check{P}_k \\ \hat{x}_k = Kz_k + (I - KC_k)\check{x}_k = \check{x}_k + K(z_k - C_k\check{x}_k) \end{cases}$$

2. 卡尔曼增益:

书中到这里就停止了，对卡尔曼增益的推导没有进行。下面是对卡尔曼增益的推导：最小化后验方差的迹对卡尔曼增益的导数：

$$min\frac{\partial Tr(\hat{P})}{K}$$

首先我们需要表示一下后验均值的误差, 这里使用了以下式子：$z_k = C_k x_k + v_k$，$\hat{x}_k = \check{x}_k + K(z_k - C_k\check{x}_k)$

$$\hat{e}_k = x_k - \hat{x}_k = (x_k - \check{x}_k) - K(z_k - C_k\check{x}_k) = (x_k - \check{x}_k) - K(C_k x_k + v_k - C_k\check{x}_k)$$

$$= (I - KC_k)(x_k - \check{x}_k) - Kv_k$$

我们可以将其代入后验方差中：

$$\hat{P}_k = E[\hat{e}_k\hat{e}_k^T] = E[(I - KC_k)(x_k - \check{x}_k)(x_k - \check{x}_k)^T(I - KC_k)^T] + E[Kv_kv_k^TK^T]$$

$$-E[(I - KC_k)(x_k - \check{x}_k)v_k^TK^T] - E[Kv_k(x_k - \check{x}_k)^T(I - KC_k)^T]$$

由于 v_k 是我们假设的正态分布，其中均值为 0，协方差为 Q，同时注意到先验方差的定义$\check{P}_k = E[(x_k - \check{x}_k)(x_k - \check{x}_k)^T]$，因此我们可以将其代入上面的式子, 化简：

$$\hat{P}_k = (I - KC_k)\check{P}_k(I - KC_k)^T + KQK^T$$

接下来计算后验方差的迹，这里使用了迹的循环置换不变性、对加法的分配率、矩阵转置不改变迹和协方差矩阵是对称阵的性质：

$$Tr(\hat{P}_k) = Tr((I - KC_k)\check{P}_k(I - KC_k)^T) + Tr(KQK^T)$$

$$= Tr(\check{P}_k(I - KC_k)(I - KC_k)^T) + Tr(KQK^T)$$

$$= Tr(\check{P}_k - \check{P}_kC_k^TK^T - \check{P}_kKC_k + \check{P}_kKC_kC_k^TK^T) + Tr(KQK^T)$$

$$= Tr(\check{P}_k) - Tr(\check{P}_k^TC_k^TK^T) - Tr(\check{P}_kKC_k) + Tr(\check{P}_kKC_kC_k^TK^T) + Tr(KQK^T)$$

$$= Tr(\check{P}_k) - 2Tr(\check{P}_kKC_k) + Tr(\check{P}_kKC_kC_k^TK^T) + Tr(KQK^T)$$

接下来对后验方差的迹进行求导：

$$\frac{\partial Tr(\hat{P}_k)}{\partial K} = -2\check{P}_k C_k + 2\check{P}_k K C_k C_k^T + Q K^T$$

令导数为 0，得到：

$$K = \check{P}_k C_k^T (Q + K C_k \check{P}_k C_k^T)^{-1}$$

至此我们完成了对卡尔曼增益的推导。