现代数字信号处理

第六讲：自适应滤波（卡尔曼滤波部分）

李喆

联系地址：电子楼 357 联系邮箱：lizhe@ suda.edu.cn

苏州大学 电子信息学院

§ 6.5.1 LMS滤波器回顾

考虑一般的系统辨识设置

$$d(k) = \mathbf{x}^H(k)\mathbf{w}^o(k) + n(k)$$

目标:根据受**观测噪声**$n(k)$干扰的**期望信号（Signal of Interest, SOI）**$d(k)$估计**未知确定**系统的参数向量$\mathbf{w}^o(k)$（**最优权值向量**）。其中$\mathbf{x}(k)$表示零均值输入向量，$n(k)$表示零均值白高斯过程，方差为$\sigma_n^2 = E\{n^2(k)\}$。该线性模型同样出现在卡尔曼滤波和机器学习中，$d(k)$,$\mathbf{x}(k)$和$\mathbf{w}^o(k)$在不同领域内对应的术语不同。

领域	$d(k)$	$\mathbf{x}(k)$	$\mathbf{w}^o(k)$
自适应滤波	期望信号	输入回归量	最优权值
卡尔曼滤波	观测值	测量值	状态向量
机器学习	目标	特征	假设参数

下面我们从确定性和时不变的最优权值向量$\mathbf{w}^o(k) = \mathbf{w}^o$开始，回顾LMS算法的推导过程，给出卡尔曼滤波器。再推广到随机和时变情况下的一般意义上的卡尔曼滤波器。

§ 6.5.1 LMS滤波器回顾

LMS需要根据现有的权值向量估计值$\mathbf{w}(k+1)$，期望$d(k)$和输入信号$\mathbf{x}(k)$，递归地估计出最优参数向量$\mathbf{w}^o$。即$\mathbf{w}(k+1)=f(\mathbf{w}(k),d(k),\mathbf{x}(k))$。三者通过输出误差相关联：

$$e(k)=d(k)-\mathbf{x}^H(k)\mathbf{w}(k)$$

定义MSE准则（**输出误差功率**）为

$$\mathrm{MSE}=\xi(k)\triangleq E\{|e(k)|^2\}.$$

LMS算法采用随机梯度下降法，通过最优权值向量$\mathbf{w}^o$以$\mathbf{w}(k+1)=\mathbf{w}(k)-\mu(k)\nabla_{\mathbf{w}}\xi(k)$形式迭代近似最小化式MSE。用瞬时功率代替平均功率，$E\{|e(k)|^2\}\approx|e(k)|^2$，LMS的更新方程为

$$\mathbf{w}(k+1)=\mathbf{w}(k)+\Delta\mathbf{w}(k)=\mathbf{w}(k)+\mu(k)\mathbf{x}(k)e(k)$$

§ 6.5.1 LMS滤波器回顾

LMS的权值误差向量

$$\tilde{\mathbf{w}}(k) \triangleq \mathbf{w}^o - \mathbf{w}(k),$$

其对输出误差的贡献为

$$e(k) = \mathbf{x}^H(k)\tilde{\mathbf{w}}(k) + n(k).$$

为了与Kalman滤波的条件一致，将$\mathbf{x}(k)$视为一个**确定性**信号，并假设噪声$n(k)$和所有其他变量都是统计独立的。那么MSE可以表示为

$$\begin{aligned}\xi(k) = E\left\{\left|\mathbf{x}^H(k)\tilde{\mathbf{w}}(k) + n(k)\right|^2\right\} &= \mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k) + \sigma_n^2 \\ &\triangleq \xi_{\mathrm{ex}}(k) + \xi_{\min}\end{aligned}$$

其中$\mathbf{P}(k) \triangleq E\{\tilde{\mathbf{w}}(k)\tilde{\mathbf{w}}^H(k)\}$是误差协方差矩阵。因此，对于迭代的每一步$k$，MSE$\xi(k)$都由两项组成：时变的**超量MSE（EMSE）**$\xi_{\mathrm{ex}}(k)$和观测噪声功率$\xi_{\min} = \sigma_n^2$。

§ 6.5.1 LMS滤波器回顾

评估辨识最优系统参数$\mathbf{w}^o$时性能的另一指标是**均方偏差（MSD）**

$$\mathrm{MSD} = J(k) \triangleq E\{\|\tilde{\mathbf{w}}(k)\|^2\} = E\{\tilde{\mathbf{w}}^H(k)\tilde{\mathbf{w}}(k)\} = \mathrm{Tr}\{\mathbf{P}(k)\}.$$

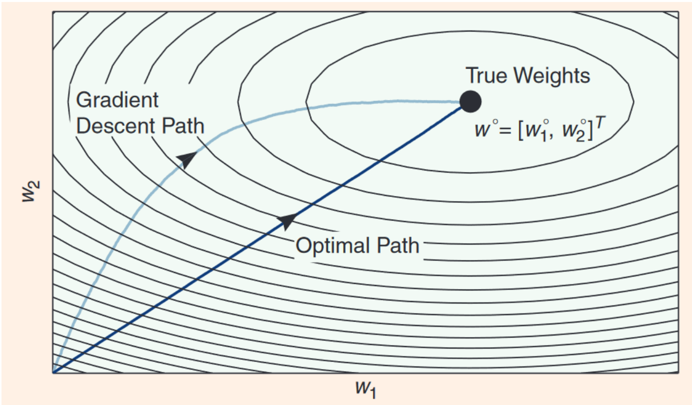
MSD表示权值误差向量的功率，通过权值误差协方差矩阵$\mathbf{P}(k) = E\{\tilde{\mathbf{w}}(k)\tilde{\mathbf{w}}^H(k)\}$观察到MSD与MSE相关，因此最小化MSD也能对应到最小化MSE。

§ 6.5.2 从LMS到卡尔曼滤波器

(1) 随机梯度算法的最优学习增益

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \Delta\mathbf{w}(k) = \mathbf{w}(k) + \mu(k)\mathbf{x}(k)e(k)$$

对于LMS的更新方程，参数$\mu(k)$是一个可能随时间变化的正步长，用于控制算法的自适应步长的大小。由于权值更新$\Delta\mathbf{w}(k) = \mu(k)\mathbf{x}(k)e(k)$和输入信号向量$\mathbf{x}(k)$方向相同，这使得LMS对数据中的离群值和噪声十分敏感。



梯度下降会执行局部最优步骤，但是却无法沿着全局最优的最短路径找到解$\mathbf{w}^o$。因此，必须同时控制自适应步骤的**方向**和**幅度**才能使算法遵循最短的最优路径，达到误差曲面的全局最小值$\xi(\mathbf{w}^o)$。

§ 6.5.2 从LMS到卡尔曼滤波器

卡尔曼滤波器的**第一步**是通过用一个**正定的学习增益矩阵**$\mathbf{G}(k)$代替标量步长$\mu(k)$引入更多自由度，从而控制梯度下降适应的大小和方向，并遵循上图中的**最优路径**。因此，更新方程变为

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{G}(k)\mathbf{x}(k)e(k)$$

与通过$\partial\xi(k)/\partial\mu(k)$最小化MSE的标准梯度自适应步长方法不同，卡尔曼的目标是在直接**最小化MSD**的基础上，在LMS中引入一个**最优步长**（学习增益）。考虑一般的迭代权值估计器

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{g}(k)e(k)$$

其中增益向量为

$$\mathbf{g}(k) \triangleq \begin{cases} \mu(k)\mathbf{x}(k), & \text{对于传统的LMS,} \\ \mathbf{G}(k)\mathbf{x}(k), & \text{对于一般的LMS.} \end{cases}$$

§ 6.5.2 从LMS到卡尔曼滤波器

为了**最小化MSD**，首先建立一般LMS的权值误差向量迭代式，在一般LMS的权值误差向量迭代式两边同时减去$\mathbf{w}^o$并用$e(k)=\mathbf{x}^H(k)\tilde{\mathbf{w}}(k)+n(k)$代替输出误差，得到

$$\tilde{\mathbf{w}}(k+1)=\tilde{\mathbf{w}}(k)-\mathbf{g}(k)\mathbf{x}^H(k)\tilde{\mathbf{w}}(k)-\mathbf{g}(k)n(k).$$

此时，权值误差协方差矩阵$\mathbf{P}(k)$的迭代式是将式上式**两边分别乘各自共轭转置并取统计期望**$E\{\cdot\}$得到

$$\begin{aligned}\mathbf{P}(k+1)&=E\{\tilde{\mathbf{w}}(k+1)\tilde{\mathbf{w}}^H(k+1)\}\\&=\mathbf{P}(k)-\left(\mathbf{P}(k)\mathbf{x}(k)\mathbf{g}^H(k)+\mathbf{g}(k)\mathbf{x}^H(k)\mathbf{P}(k)\right)\\&\quad+\mathbf{g}(k)\mathbf{g}^H(k)\left(\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k)+\sigma_n^2\right)\end{aligned}$$

根据矩阵迹的性质得到

$$\mathrm{Tr}\{\mathbf{P}(k)\mathbf{x}(k)\mathbf{g}^H(k)\}=\mathbf{g}^H(k)\mathbf{P}(k)\mathbf{x}(k)$$

且$\mathrm{Tr}\{\mathbf{g}(k)\mathbf{g}^H(k)\}=\|\mathbf{g}(k)\|^2$，那么MSD $J(k)=\mathrm{Tr}\{\mathbf{P}(k)\}$的迭代式为

$$\begin{aligned}J(k+1)=&J(k)-\mathbf{g}^H(k)\mathbf{P}(k)\mathbf{x}(k)-\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{g}(k)\\&+\|\mathbf{g}(k)\|^2\left(\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k)+\sigma_n^2\right)\end{aligned}$$

§ 6.5.2 从LMS到卡尔曼滤波器

(2) LMS最优标量步长(与NLMS的关系)

LMS的最优步长是为了实现$e(k+1|k) = d(k) - \mathbf{x}^H(k)\mathbf{w}(k) = 0$，其中使用更新的权值向量$\mathbf{w}(k)$和当前输入$\mathbf{x}(k)$得到后验误差$e(k+1|k)$。这种方法也叫NLMS，即

$$\text{NLMS}: \quad \mathbf{w}(k+1) = \mathbf{w}(k) + \frac{1}{\|\mathbf{x}(k)\|^2}\mathbf{x}(k)e(k)$$

有效的LMS型步长$\mu(k) = \frac{1}{\|\mathbf{x}(k)\|^2}$是时变的，实际中往往引入小的正步长$\rho(k)$使算法稳定，即$\mu(k) = \frac{\rho(k)}{\|\mathbf{x}(k)\|^2}$。如果使MSD最小化，得到LMS的最优标量步长，首先将增益$\mathbf{g}(k) = \mu(k)\mathbf{x}(k)$代入$J(k)$迭代式，给出MSD的迭代式

$$J(k+1) = J(k) - 2\mu(k)\underbrace{\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k)}_{\xi_{\text{ex}}(k)}$$

$$+ \mu^2(k)\|\mathbf{x}(k)\|^2\,(\underbrace{\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k) + \sigma_n^2}_{\xi(k)})$$

§ 6.5.2 从LMS到卡尔曼滤波器

通过$\frac{\partial J(k)}{\partial \mu(k)} = 0$解得上式中的$\mu(k)$来获得最小化MSD的最优步长，

$$\mu(k) = \frac{1}{\parallel \mathbf{x}(k) \parallel^2} \frac{\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k)}{(\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k) + \sigma_n^2)} = \underbrace{\frac{1}{\|\mathbf{x}(k)\|^2}}_{\text{normalization}} \underbrace{\frac{\xi_{\text{ex}}(k)}{\xi(k)}}_{\text{correction}}$$

除了NLMS型归一化因子$\frac{1}{\|\mathbf{x}(k)\|^2}$，最优LMS步长包含修正项$\frac{\xi_{\text{ex}}(k)}{\xi(k)} < 1$，表示EMSE$\xi_{\text{ex}}(k)$与MSE$\xi(k)$的比率。与真实系统权值的较大偏差导致了较大的$\frac{\xi_{\text{ex}}(k)}{\xi(k)}$和较快的权值调整（与较小$\frac{\xi_{\text{ex}}(k)}{\xi(k)}$引起的较慢调整相比）。

§ 6.5.2 从LMS到卡尔曼滤波器

(3) 确定性状态下的Kalman滤波器

在**第(2)**步，最优LMS步长旨在每一个瞬时时刻都最小化MSD，它**只能控制梯度步长的大小**。为了同时控制梯度下降的大小和方向，回到MSD的表达式。通过$\frac{\partial J(k)}{\partial \mathbf{g}(k)}=0$解得上述MSD的$\mathbf{g}(k)$，从而得到最优学习增益向量$\mathbf{g}(k)$，即

$$\mathbf{g}(k)=\frac{\mathbf{P}(k)}{\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k)+\sigma_n^2}\mathbf{x}(k)=\frac{\mathbf{P}(k)}{\xi(k)}\mathbf{x}(k)=\mathbf{G}(k)\mathbf{x}(k)$$

这个最优增益向量$\mathbf{g}(k)$就是**卡尔曼增益**，而**增益矩阵**$\mathbf{G}(k)$表示权值误差协方差矩阵$\mathbf{P}(k)$与MSE$\xi(k)$之间的比值。此时权误差协方差矩阵的$\mathbf{P}(k)$的更新方程变为

$$\begin{aligned}\mathbf{P}(k+1)=\mathbf{P}(k)&-\left(\mathbf{P}(k)\mathbf{x}(k)\mathbf{g}^H(k)+\mathbf{g}(k)\mathbf{x}^H(k)\mathbf{P}(k)\right)\\&+\mathbf{g}(k)\mathbf{g}^H(k)\left(\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k)+\sigma_n^2\right)\\=\mathbf{P}(k)&-\mathbf{g}(k)\mathbf{x}^H(k)\mathbf{P}(k)\end{aligned}$$

§ 6.5.2 从LMS到卡尔曼滤波器

通过更新$\mathbf{P}(k)$就可以得到卡尔曼滤波器，即：

1) 计算最优学习增益（**卡尔曼增益**）：

$$\mathbf{g}(k) = \mathbf{P}(k)\mathbf{x}(k)/(\mathbf{x}^H(k)\mathbf{P}(k)\mathbf{x}(k) + \sigma_n^2)$$

2) 更新权值向量：

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{g}(k)(d(k) - \mathbf{x}^H(k)\mathbf{w}(k))$$

3) 更新权值误差协方差矩阵：

$$\mathbf{P}(k+1) = \mathbf{P}(k) - \mathbf{g}(k)\mathbf{x}^H(k)\mathbf{P}(k)$$

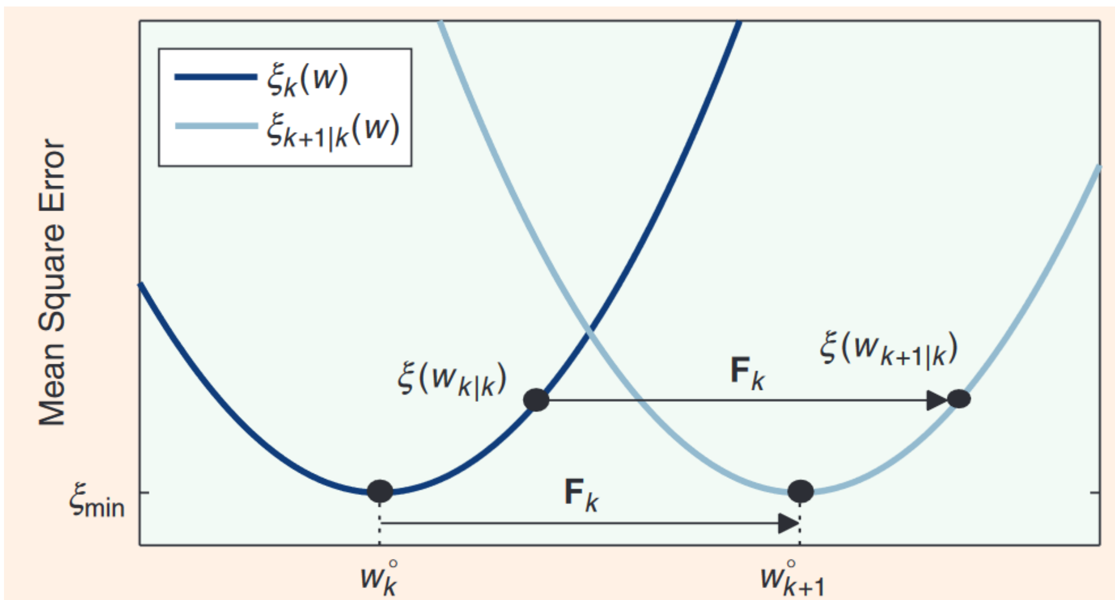
注意：对于$\sigma_n^2 = 1$，上述算法的更新公式与**递归最小二乘（RLS）算法**一致。

§ 6.5.2 从LMS到卡尔曼滤波器

(4) 一般的卡尔曼滤波

下面考虑权值向量$\mathbf{w}^o(k)$时变的情况，从最优LMS导出一般卡尔曼滤波的算法。考虑真实的权值向量$\mathbf{w}^o(k)$的演变由已知的**状态转换矩阵$\mathbf{F}(k)$**控制，而状态转换中的不确定性由协方差为$\mathbf{Q}_s = E\{\mathbf{q}(k)\mathbf{q}^H(k)\}$的噪声向量$\mathbf{q}(k)$表示，其与观测噪声$n(k)$**不相关**。那么

$$\mathbf{w}^o(k+1) = \mathbf{F}(k)\mathbf{w}^o(k) + \mathbf{q}(k), \qquad \mathbf{q}(k) \sim \mathcal{N}(0, \mathbf{Q}_s)$$



§ 6.5.2 从LMS到卡尔曼滤波器

系统辨识模型仍为：

$$d(k) = \mathbf{x}^H(k)\mathbf{w}^o(k) + n(k), \qquad n(k) \sim \mathcal{N}(0, \sigma_n^2)$$

最优权值向量的更新分两步

1）以当前状态值$\mathbf{w}(k|k)$预测下一状态$\mathbf{w}(k+1|k)$

$$\mathbf{w}(k+1|k) = \mathbf{F}(k)\mathbf{w}(k|k),$$

2）以类似LMS方式更新下一状态估计值$\mathbf{w}(k+1|k+1)$

$$\mathbf{w}(k+1|k+1) = \mathbf{w}(k+1|k) + \mathbf{g}(k)[d(k) - \mathbf{x}^H(k)\mathbf{w}(k+1|k)],$$

其中 $\mathbf{g}(k)$是**卡尔曼增益**。

§ 6.5.2 从LMS到卡尔曼滤波器

所有相关量的更新步骤（用$k|k$表示）和预测步骤（用$k+1|k$表示）定义如下

$$\begin{aligned}
\tilde{\mathbf{w}}(k|k) &\triangleq \mathbf{w}^o(k) - \mathbf{w}(k|k), \\
\mathbf{P}(k|k) &\triangleq E\{\tilde{\mathbf{w}}(k|k)\tilde{\mathbf{w}}^H(k|k)\}, \\
\tilde{\mathbf{w}}(k+1|k) &\triangleq \mathbf{w}^o(k+1) - \mathbf{w}(k+1|k) = \mathbf{F}(k)\tilde{\mathbf{w}}(k|k) + \mathbf{q}(k), \\
\mathbf{P}(k+1|k) &\triangleq E\{\tilde{\mathbf{w}}(k+1|k)\tilde{\mathbf{w}}^H(k+1|k)\} \\
&= \mathbf{F}(k)\mathbf{P}(k|k)\mathbf{F}^H(k) + \mathbf{Q}_s.
\end{aligned}$$

因此，权值误差向量的更新公式为

$$\tilde{\mathbf{w}}(k+1|k+1) = \tilde{\mathbf{w}}(k+1|k) - \mathbf{g}(k)\mathbf{x}^H(k)\tilde{\mathbf{w}}(k+1|k) - \mathbf{g}(k)n(k),$$

§ 6.5.2 从LMS到卡尔曼滤波器

从而得到权值误差协方差的迭代式为

$$\begin{aligned}\mathbf{P}(k+1|k+1) &\triangleq E\{\tilde{\mathbf{w}}(k+1|k+1)\tilde{\mathbf{w}}^H(k+1|k+1)\} \\ &= \mathbf{P}(k+1|k) - \left[\mathbf{P}(k+1|k)\mathbf{x}(k)\mathbf{g}^H(k) + \mathbf{g}(k)\mathbf{x}^H(k)\mathbf{P}(k+1|k)\right] \\ &\quad + \mathbf{g}(k)\mathbf{g}^H(k)\left[\mathbf{x}^H(k)\mathbf{P}(k+1|k)\mathbf{x}(k) + \sigma_n^2\right]\end{aligned}$$

最后，最小化MSD $J(k|k) = \mathrm{Tr}\{\mathbf{P}(k+1|k+1)\}$，得到卡尔曼增益$\mathbf{g}(k)$的表达式为

$$\mathbf{g}(k) = \frac{\mathbf{P}(k+1|k)}{\mathbf{x}^H(k)\mathbf{P}(k+1|k)\mathbf{x}(k) + \sigma_n^2}\mathbf{x}(k) = \mathbf{G}(k)\mathbf{x}(k)$$

其与最优LMS增益相同。

§ 6.5.2 从LMS到卡尔曼滤波器

一般卡尔曼滤波器的算法步骤：对于每个时刻$k > 0$，基于测量值$\{d(k), \mathbf{x}(k)\}$有

1) 预测下一个（后验）权值向量（状态）：

$$\mathbf{w}(k+1|k) = \mathbf{F}(k)\mathbf{w}(k|k)$$

2) 预测权值误差协方差矩阵：

$$\mathbf{P}(k+1|k) = \mathbf{F}(k)\mathbf{P}(k|k)\mathbf{F}^H(k) + \mathbf{Q}_s$$

3) 计算最优学习增益（卡尔曼增益）：

$$\mathbf{g}(k) = \mathbf{P}(k+1|k)\mathbf{x}(k)/(\mathbf{x}^H(k)\mathbf{P}(k+1|k)\mathbf{x}(k) + \sigma_n^2)$$

4) 更新权值向量估计：

$$\mathbf{w}(k+1|k+1) = \mathbf{w}(k+1|k) + \mathbf{g}(k)(d(k) - \mathbf{x}^H(k)\mathbf{w}(k+1|k))$$

5) 更新权值误差协方差矩阵：

$$\mathbf{P}(k+1|k+1) = \mathbf{P}(k+1|k) - \mathbf{g}(k)\mathbf{x}^H(k)\mathbf{P}(k+1|k)$$

§ 6.5.3 基于贝叶斯理论的离散卡尔曼滤波

贝叶斯理论：$P(X|Z)P(Z) = P(Z|X)P(X)$，其中$P(X)$为**先验分布**，$P(Z|X)$为**似然函数**。$P(X|Z)$为**后验分布**，表示结合观测信息后的状态概率更新。卡尔曼滤波可看作是**贝叶斯更新在线性高斯动态系统中的具体实现**，包括**先验更新（预测阶段）**和**后验更新（修正阶段）**。Kalman滤波器的**两个方程**：

$$\begin{cases} \textbf{矢量状态方程:} & \boldsymbol{x}(n) = \boldsymbol{A}(n-1)\boldsymbol{x}(n-1) + \boldsymbol{q}(n) \\ \textbf{矢量观测方程:} & \boldsymbol{y}(n) = \boldsymbol{C}(n)\boldsymbol{x}(n) + \boldsymbol{v}(n) \end{cases}$$

定义: 推导中用 $\hat{\boldsymbol{x}}(n/n)$ 表示给定 $i = 1, 2, \ldots, n$ 的观测 $\boldsymbol{y}(i)$ 而在时刻 n 获得的 $\boldsymbol{x}(n)$ 的最佳线性估计，用 $\hat{\boldsymbol{x}}(n/n-1)$ 表示给定 $i = 1, 2, \ldots, n-1$ 的观测 $\boldsymbol{y}(i)$ 时 $\boldsymbol{x}(n)$ 的最佳估计。

用 $\boldsymbol{e}(n/n)$ 和 $\boldsymbol{e}(n/n-1)$ 表示相应的**状态估计误差:**

$$\boldsymbol{e}(n/n) = \boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n/n)$$

$$\boldsymbol{e}(n/n-1) = \boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n/n-1)$$

误差协方差阵定义为:

$$\begin{cases} \boldsymbol{P}(n/n) = E\left\{\boldsymbol{e}(n/n)\boldsymbol{e}^H(n/n)\right\} \\ \boldsymbol{P}(n/n-1) = E\left\{\boldsymbol{e}(n/n-1)\boldsymbol{e}^H(n/n-1)\right\} \end{cases}$$

整体求解过程—迭代实现

- 假设给定 $\boldsymbol{x}(0)$ 的一个估计 $\hat{\boldsymbol{x}}(0/0)$，相应于这一估计的误差协方差阵 $\boldsymbol{P}(0/0)$ 也已知.

- 当观测值 $\boldsymbol{y}(1)$ 到来时, 要修正 $\hat{\boldsymbol{x}}(0/0)$ 以获得 n=1 时刻的状态估计 $\hat{\boldsymbol{x}}(1/1)$，并使如下均方误差最小:

$$\xi(1)=E\left\{\|\boldsymbol{e}(1/1)\|^2\right\}=\operatorname{tr}\{\boldsymbol{P}(1/1)\}=\sum_{i=0}^{p-1}E\left\{\left|e_i(1/1)\right|^2\right\}$$

- 确定 $\hat{\boldsymbol{x}}(1/1)$ 后, 就可得到误差协方差阵 $\boldsymbol{P}(1/1)$，再随着观测值 $y(2)$ 的到来重复以上估计过程。⋯⋯到第 n 步; 直到稳态收敛, 并保持、跟踪.

每次迭代求解

- 这样对每个$n>0$, 给定 $\hat{\boldsymbol{x}}(n-1/n-1)$ 和 $\boldsymbol{P}(n-1/n-1)$,当新观测值 $\boldsymbol{y}(n)$ 到来时, 要求解状态矢量 $\boldsymbol{x}(n)$ 新的最小均方误差估计 $\hat{\boldsymbol{x}}(n/n)$ 。
- **求解该问题可分为两步:**
- **第一步是预测估计**, 由 $\hat{\boldsymbol{x}}(n-1/n-1)$ 获得 $\hat{\boldsymbol{x}}(n/n-1)$,它是无 $y(n)$ 观测值时 $x(n)$ 的最佳估计;
- **第二步是滤波估计**,由 $y(n)$ 和 $\hat{\boldsymbol{x}}(n/n-1)$ 的最优组合来估计 $x(n)$ 。

§ 6.5.3 基于贝叶斯理论的离散卡尔曼滤波

第一步: 无新观测值, 只能利用 $\boldsymbol{x}(n)$ 状态方程:

$\boldsymbol{x}(n)=\boldsymbol{A}(n-1)\boldsymbol{x}(n-1)+\boldsymbol{q}(n)$, 得 $\boldsymbol{x}(n)$ 的预测估计:

$$\widehat{\boldsymbol{x}}(n/n-1)=\boldsymbol{A}(n-1)\hat{\boldsymbol{x}}(n-1/n-1)$$

$$\begin{aligned}\boldsymbol{e}(n/n-1)&=\boldsymbol{x}(n)-\widehat{\boldsymbol{x}}(n/n-1)\end{aligned}$$

误差为:

$$\begin{aligned}&=\boldsymbol{A}(n-1)\boldsymbol{x}(n-1)+\boldsymbol{q}(n)-\boldsymbol{A}(n-1)\widehat{\boldsymbol{x}}(n-1/n-1)\\&=\boldsymbol{A}(n-1)\boldsymbol{e}(n-1/n-1)+\boldsymbol{q}(n)\end{aligned}$$

$\longrightarrow$ 因 $\boldsymbol{q}(n)$ 均值为零, 若 $\hat{\boldsymbol{x}}(n-1/n-1)$ 是无偏估计, 即

$E\{\boldsymbol{e}(n-1/n-1)\}=0$, 则 $\hat{\boldsymbol{x}}(n/n-1)$ 也是 $\boldsymbol{x}(n)$ 的无偏估计, 即:

$E\{\boldsymbol{e}(n/n-1)\}=0$.

最后, 因估计误差 $\boldsymbol{e}(n-1/n-1)$ 与 $\boldsymbol{q}(n)$ 不相关， $\boldsymbol{q}(n)$ 是白噪声序列, 因此:

$\boldsymbol{P}(n/n-1)=\boldsymbol{A}(n-1)\boldsymbol{P}(n-1/n-1)\boldsymbol{A}^H(n-1)+\boldsymbol{Q}_q(n)$ 其中 $\boldsymbol{Q}_q(n)$ 是噪声 $\boldsymbol{q}(n)$ 的协方差阵。 **以上即卡尔曼滤波第一步。**

§ 6.5.3 基于贝叶斯理论的离散卡尔曼滤波

第二步: 要将新观测 $\boldsymbol{y}(n)$ 与 $\hat{\boldsymbol{x}}(n/n-1)$ 结合**得到新估计:**

$$\widehat{\boldsymbol{x}}(n/n) = \boldsymbol{K}'(n)\widehat{\boldsymbol{x}}(n/n-1) + \boldsymbol{K}(n)\boldsymbol{y}(n)$$

其中 $\boldsymbol{K}(n)$ 和 $\boldsymbol{K}'(n)$ 是要确定的矩阵。

1) 首先由 $\boldsymbol{x}(n)$ 估计的无偏性得 $\boldsymbol{K}(n)$ 和 $\boldsymbol{K}'(n)$ 的关系

$\hat{\boldsymbol{x}}(n/n)$ 无偏 $\Leftrightarrow E\{\boldsymbol{e}(n/n)\} = 0$. 由(I) $\rightarrow$ 用 $\boldsymbol{e}(n/n-1)$ 表达的 $\boldsymbol{e}(n/n)$:

$$\begin{aligned}
\boldsymbol{e}(n/n) &= \boldsymbol{x}(n) - \boldsymbol{K}'(n)\hat{\boldsymbol{x}}(n/n-1) - \boldsymbol{K}(n)\boldsymbol{y}(n) \\
&= \boldsymbol{x}(n) - \boldsymbol{K}'(n)[\boldsymbol{x}(n) - \boldsymbol{e}(n/n-1)] - \boldsymbol{K}(n)[\boldsymbol{C}(n)\boldsymbol{x}(n) + \boldsymbol{v}(n)] \\
&= \left[\boldsymbol{I} - \boldsymbol{K}'(n) - \boldsymbol{K}(n)\boldsymbol{C}(n)\right]\boldsymbol{x}(n) + \boldsymbol{K}'(n)\boldsymbol{e}(n/n-1) - \boldsymbol{K}(n)\boldsymbol{v}(n)
\end{aligned}$$

由于 $E\{\boldsymbol{v}(n)\} = 0, E\{\boldsymbol{e}(n/n-1)\} = 0$，若要上式的 $E\{\boldsymbol{e}(n/n)\} = 0$，

必须满足: $\boldsymbol{K}'(n) = \boldsymbol{I} - \boldsymbol{K}(n)\boldsymbol{C}(n)$

代入(I)得: $\widehat{\boldsymbol{x}}(n/n) = [\boldsymbol{I} - \boldsymbol{K}(n)\boldsymbol{C}(n)]\widehat{\boldsymbol{x}}(n/n-1) + \boldsymbol{K}(n)\boldsymbol{y}(n)$

重组: $\widehat{\boldsymbol{x}}(n/n) = \widehat{\boldsymbol{x}}(n/n-1) + \boldsymbol{K}(n)[\boldsymbol{y}(n) - \boldsymbol{C}(n)\hat{\boldsymbol{x}}(n/n-1)]$

2)再由 $\min E\left\{\|\boldsymbol{e}(n/n)\|^2\right\} = \min \operatorname{tr}\{\boldsymbol{P}(n/n)\}$ **准则得最优** $\boldsymbol{K}(n)$:误差准则: $\xi(n) = \operatorname{tr}\{\boldsymbol{P}(n/n)\}$, 用 $d\xi/d\boldsymbol{K}(n) = 0$ 求解 $\boldsymbol{K}(n)$ 。

误差:

$$\begin{aligned}\boldsymbol{e}(n/n) &= \boldsymbol{K}'(n)\boldsymbol{e}(n/n-1) - \boldsymbol{K}(n)\boldsymbol{v}(n)\\ &= [\boldsymbol{I} - \boldsymbol{K}(n)\boldsymbol{C}(n)]\boldsymbol{e}(n/n-1) - \boldsymbol{K}(n)\boldsymbol{v}(n)\end{aligned}$$

因 $\boldsymbol{v}(n)$ 与 $\boldsymbol{q}(n)$ (即 $\boldsymbol{x}(n)$) 不相关, 也与 $\hat{\boldsymbol{x}}(n/n-1)$ 不相关,而 $\boldsymbol{e}(n/n-1) = \boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n/n-1)$, 故上式中 $\boldsymbol{v}(n)$ 与 $\boldsymbol{e}(n/n-1)$ 也不相关。 $\longrightarrow$

$$\begin{aligned}&\boldsymbol{P}(n/n) = E\left\{\boldsymbol{e}(n/n)\boldsymbol{e}^H(n/n)\right\}\\ &= [\boldsymbol{I} - \boldsymbol{K}(n)\boldsymbol{C}(n)]\boldsymbol{P}(n/n-1)[\boldsymbol{I} - \boldsymbol{K}(n)\boldsymbol{C}(n)]^H + \boldsymbol{K}(n)\boldsymbol{Q}_v(n)\boldsymbol{K}^H(n)\end{aligned}$$

再给一些公式: $\frac{d}{d\boldsymbol{K}}\operatorname{tr}\left(\boldsymbol{K}\boldsymbol{A}\boldsymbol{K}^H\right) = 2\boldsymbol{K}\boldsymbol{A}, \frac{d}{d\boldsymbol{K}}\operatorname{tr}(\boldsymbol{K}\boldsymbol{A}) = \boldsymbol{A}^H$

§ 6.5.3 基于贝叶斯理论的离散卡尔曼滤波

最后得式(II):

$\frac{d}{d\boldsymbol{K}}\,\mathrm{tr}[\boldsymbol{P}(n/n)] = -2[\boldsymbol{I}-\boldsymbol{K}(n)\boldsymbol{C}(n)]\boldsymbol{P}(n/n-1)\boldsymbol{C}^H(n)+2\boldsymbol{K}(n)\boldsymbol{Q}_v(n) = 0$

求得:

$$\boldsymbol{K}(n) = \boldsymbol{P}(n/n-1)\boldsymbol{C}^H(n)\left[\boldsymbol{C}(n)\boldsymbol{P}(n/n-1)\boldsymbol{C}^H(n)+\boldsymbol{Q}_v(n)\right]^{-1}$$

导出卡尔曼增益矢量后, 可用它再简化前面误差协方差阵 $\boldsymbol{P}(n/n)$ 的表达式, 得:

$$\begin{aligned}\boldsymbol{P}(n/n) &= [\boldsymbol{I}-\boldsymbol{K}(n)\boldsymbol{C}(n)]\boldsymbol{P}(n/n-1) \\ &\quad - \left\{[\boldsymbol{I}-\boldsymbol{K}(n)\boldsymbol{C}(n)]\boldsymbol{P}(n/n-1)\boldsymbol{C}^H(n)+\boldsymbol{K}(n)\boldsymbol{Q}_v(n)\right\}\boldsymbol{K}^H(n)\end{aligned}$$

上式第2项应为零, $\boldsymbol{P}(n/n)$ 可简化为:

$$\boldsymbol{P}(n/n) = [\boldsymbol{I}-\boldsymbol{K}(n)\boldsymbol{C}(n)]\boldsymbol{P}(n/n-1)$$

至此已得估计 $\boldsymbol{x}(n)$ 的卡尔曼滤波递归式.

剩下问题是 $n=0$ 时的初始化问题。

因初始状态值未知, 在没有 $n=0$ 时刻观测数据时, 可选择初始状态估计为: $\hat{\boldsymbol{x}}(0/0)=E\{\boldsymbol{x}(0)\}$

而对误差协方差阵, 初始值取为:

$$\boldsymbol{P}(0/0)=E\left\{\boldsymbol{x}(0)\boldsymbol{x}^{H}(0)\right\}$$

这种选择使得 $\hat{\boldsymbol{x}}(0/0)$ 是 $\boldsymbol{x}(0)$ 的无偏估计，从而保证对所有的 n, $\hat{\boldsymbol{x}}(n/n)$ 是无偏的(注意推导卡尔曼迭代方程时, 约束了 $\hat{\boldsymbol{x}}(n/n)$ 是无偏的)。

另有一点要注意: 卡尔曼增益 $\boldsymbol{K}(n)$ 和误差协方差阵 $\boldsymbol{P}(n/n)$ 的递归估计与数据 $\boldsymbol{x}(n)$ 无关, 故可离线预先计算出来,用于任意数据滤波.

总结：离散卡尔曼滤波

已知条件: **状态方程** $\boldsymbol{x}(n) = \boldsymbol{A}(n-1)\boldsymbol{x}(n-1) + \boldsymbol{q}(n)$

观测方程 $\boldsymbol{y}(n) = \boldsymbol{C}(n)\boldsymbol{x}(n) + \boldsymbol{v}(n)$

初始化: $\hat{\boldsymbol{x}}(0/0) = E\{\boldsymbol{x}(0)\}$, $\boldsymbol{P}(0/0) = E\left\{\boldsymbol{x}(0)\boldsymbol{x}^H(0)\right\}$

递归估计算法:

1. **预测** $\hat{\boldsymbol{x}}(n/n-1) = \boldsymbol{A}(n-1)\hat{\boldsymbol{x}}(n-1/n-1)$
2. **MSE** $\boldsymbol{P}(n/n-1) = \boldsymbol{A}(n-1)\boldsymbol{P}(n-1/n-1)\boldsymbol{A}^H(n-1)+\boldsymbol{Q}_q(n)$
3. **增益** $\boldsymbol{K}(n)=\boldsymbol{P}(n/n-1)\boldsymbol{C}^H(n)\left[\boldsymbol{C}(n)\boldsymbol{P}(n/n-1)\boldsymbol{C}^H(n)+\boldsymbol{Q}_v(n)\right]^{-1}$
4. **修正** $\hat{\boldsymbol{x}}(n/n) = \widehat{\boldsymbol{x}}(n/n-1) + \boldsymbol{K}(n)[\boldsymbol{y}(n) - \boldsymbol{C}(n)\hat{\boldsymbol{x}}(n/n-1)]$
5. **MSE** $\boldsymbol{P}(n/n) = [\boldsymbol{I} - \boldsymbol{K}(n)\boldsymbol{C}(n)]\boldsymbol{P}(n/n-1)$

§ 6.5.3 基于贝叶斯理论的离散卡尔曼滤波

例6.1：利用卡尔曼滤波估计一个未知的常数。设已知一个未知常数 x 的噪声观测集合，已知噪声 $v(n)$ 的均值为零，方差为 σ_v^2，$v(n)$ 与 x 不相关，试用卡尔曼滤波估计该常数。

解：由于 x 不随时间 n 变化，因此**状态方程**为：$x(n)=x(n-1)$

而**观测方程**为：$$y(n)=x(n)+v(n)$$

因此，这时的 $\boldsymbol{A}(n)=1$,$\boldsymbol{C}(n)=1$, $\boldsymbol{Q}_v(n)=\sigma_v^2$, $\boldsymbol{Q}_q(n)=0$ 。又 $x(n)$ 是标量，误差协方差也是标量，可表示为：$P(n/n)=E\left\{e^2(n/n)\right\}$其中 $e(n/n)=x(n)-\hat{x}(n/n)$ 。这时的 $P(n/n-1)$ 和 $P(n-1/n-1)$ 是相等的，即：$P(n/n-1)=P(n-1/n-1)$

因此为表达简单，**我们将** $P(n/n-1)$ **和** $P(n-1/n-1)$ **都用** $P(n-1)$ **来表示**。由卡尔曼增益表达式，这时有：

$$K(n)=P(n-1)\left[P(n-1)+\sigma_v^2\right]^{-1}$$

§ 6.5.3 基于贝叶斯理论的离散卡尔曼滤波

$P(n/n)$ **迭代公就为:**

$$P(n)=[1-K(n)]P(n-1)=\left[1-\frac{P(n-1)}{P(n-1)+\sigma_v^2}\right]P(n-1)=\frac{P(n-1)\sigma_v^2}{P(n-1)+\sigma_v^2}$$

我们可以用递归的方式来求解上式的差分方程:

$$P(1)=\frac{P(0)\sigma_v^2}{P(0)+\sigma_v^2},P(2)=\frac{P(1)\sigma_v^2}{P(1)+\sigma_v^2}=\frac{P(0)\sigma_v^2}{2P(0)+\sigma_v^2},$$
$$P(3)=\frac{P(2)\sigma_v^2}{P(2)+\sigma_v^2}=\frac{P(0)\sigma_v^2}{3P(0)+\sigma_v^2},\cdots\cdots\to P(n)=\frac{P(0)\sigma_v^2}{nP(0)+\sigma_v^2}$$

将上式结合到卡尔曼增益 $K(n)$ **的计算式中，有:**

$$K(n)=\frac{P(n-1)}{P(n-1)+\sigma_v^2}=\frac{P(0)}{nP(0)+\sigma_v^2}$$

最后得离散卡尔曼滤波器的计算式:

$$\widehat{x}(n)=\widehat{x}(n-1)+\frac{P(0)}{nP(0)+\sigma_v^2}[y(n)-\widehat{x}(n-1)]$$

注意当 $n\to\infty$ **时，**$K(n)\to 0$ **，因此** $x^\wedge(n)$ **趋于稳态值。**

§ 6.5.3 基于贝叶斯理论的离散卡尔曼滤波

考虑两种较为极端的特殊情况。

第一种情况, 若 $\sigma_v^2 \to \infty$, 即观测值完全不可靠, 则卡尔曼增益趋于零, 估计公式变成: $\widehat{x}(n) = \widehat{x}(n-1)$

因此, 观测值被忽略掉, $\widehat{x}(n) = \widehat{x}(0)$, 保持为初始估计, 误差方程为 $P(0)$ 。

第二种情况, 假设 $\widehat{x}(0) = 0$,且 $P(0) \to \infty$, 这相当于没有 x 的先验信息。这时 $K(n) = 1/n$, 估计式变为:

$$\widehat{x}(n) = \widehat{x}(n-1) + (1/n)[y(n) - \widehat{x}(n-1)]$$

$$= [(n-1)/n]\widehat{x}(n-1) + (1/n)y(n)$$

这实际上是递归地实现如下的采样均值计算:

$$\widehat{x}(n) = \frac{1}{n}\sum_{k=1}^{n} y(k)$$

§ 6.5.3 基于贝叶斯理论的离散卡尔曼滤波

例6.2：设 $x(n)$ 是如下的 AR(1) 过程: $x(n)=0.8x(n-1)+w(n)$，其中 $w(n)$ 是白噪声，方差为 $\sigma_w{}^2=0.36$。又用 $y(n)=x(n)+v(n)$ 表示信号 $x(n)$ 的噪声观测，$v(n)$ 是单位方差的白噪声，与 $w(n)$ 不相关。

解: 这时的 $\boldsymbol{A}(n)=0.8$, $\boldsymbol{C}(n)=1$, $\boldsymbol{Q}_v(n)=1$, $\boldsymbol{Q}_q(n)=0.36$，**卡尔曼滤波的状态估计方程为:**

$$\widehat{x}(n)=0.8\widehat{x}(n-1)+K(n)[y(n)-0.8\widehat{x}(n-1)]$$

因状态变量是标量，故卡尔曼增益计算公式也是标量方程:

$$P(n/n-1)=(0.8)^2P(n-1/n-1)+0.36$$

$$K(n)=P(n/n-1)[P(n/n-1)+1]^{-1}$$

$$P(n/n)=[1-K(n)]P(n/n-1)$$

取 $\hat{x}(0)=E\{x(0)\}=E\{w(n)\}=0$

$$P(0/0)=E\left\{|x(0)|^2\right\}=E\left\{\frac{|w(n)|^2}{1-0.8^2}\right\}=1$$

§ 6.5.3 基于贝叶斯理论的离散卡尔曼滤波

前面几次迭代的卡尔曼增益和误差协方差如下表所示

n	P(n / n-1)	K(n)	P(n / n)
1	1.0000	0.5000	0.5000
2	0.6800	0.4048	0.4048
3	0.6190	0.3624	0.3824
4	0.6047	0.3768	0.3768
5	0.6012	0.3755	0.3755
$\vdots$	$\vdots$	$\vdots$	$\vdots$
∞	0.6000	0.3750	0.3750

可以看到，卡尔曼滤波器是**时变的**,且在 $n = 0$ 时刻就开始工作。但因处理的是**平稳过程**, 它在经历初始瞬变后, 逐渐进入稳态行为:

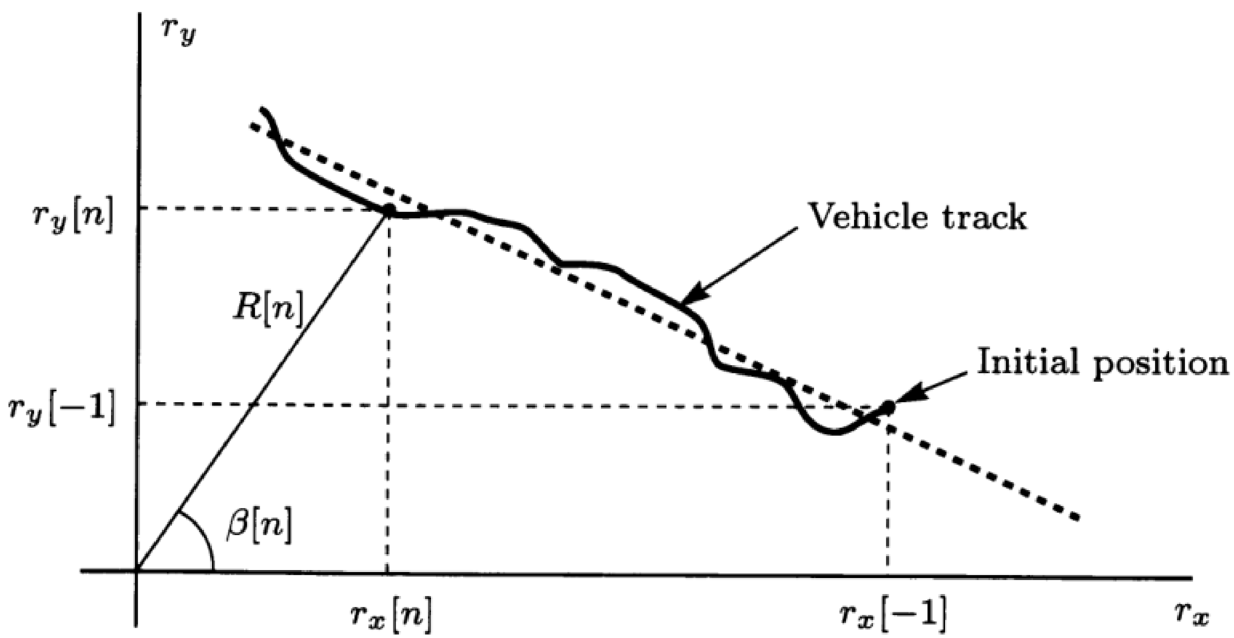
$$\widehat{x}(n) = 0.8\widehat{x}(n-1) + 0.375[y(n) - 0.8\bar{x}(n-1)]$$

最后均方误差是 $\xi = 0.375$。

§6.5.4 扩展卡尔曼滤波(EKF)

实际场景中我们经常遇到状态方程或观测方程是**非线性的**，例如运动目标的追踪问题。

$$
\begin{aligned}
R(n) &= \sqrt{r_x^2(n) + r_y^2(n)} \\
\beta(n) &= \arctan \frac{r_y(n)}{r_x(n)}
\end{aligned}
$$



§6.5.4 扩展卡尔曼滤波(EKF)

假设位置$R(n)$和方位$\beta(n)$在观测时分别受到噪声$w_R(n)$和$w_\beta(n)$的干扰，即

$$\begin{aligned}\hat{R}(n) = &\quad R(n) + w_R(n)\\ \hat{\beta}(n) = &\quad \beta(n) + w_\beta(n)\end{aligned}$$

如果存在状态方程:

$$\begin{aligned}r_x(n) &= r_x(n-1) + v_x\Delta\\ r_y(n) &= r_y(n-1) + v_y\Delta\end{aligned}$$

那么位置$R(n)$需要表示为

$$R(n) = \sqrt{R^2(n-1) + 2R(n-1)\Delta(v_x\cos\beta(n-1) + v_y\sin\beta(n-1)) + (v_x^2+v_y^2)\Delta^2}$$

定义$\boldsymbol{y}(n) = [R(n), \beta(n)]^T$和 $\boldsymbol{x}(n) = [r_x(n), r_y(n)]^T$，则$\boldsymbol{y}(n)$和$\boldsymbol{x}(n)$之间**不存在**线性观测方程

$$\boldsymbol{y}(n) = \boldsymbol{C}(n)\boldsymbol{x}(n) + \boldsymbol{v}(n)$$

§6.5.4 扩展卡尔曼滤波(EKF)

扩展卡尔曼滤波(EKF)的状态方程和观测方程:

$$\begin{cases} \boldsymbol{x}(n) = \boldsymbol{a}\{\boldsymbol{x}(n-1)\} + \boldsymbol{v}(n) \\ \boldsymbol{y}(n) = \boldsymbol{C}\{\boldsymbol{x}(n)\} + \boldsymbol{v}(n) \end{cases}$$

求解方法: 对状态转移函数$\boldsymbol{a}\{\cdot\}$和观测函数$\boldsymbol{C}\{\cdot\}$使用一阶泰勒展开得到近似值:

$$\begin{aligned} \boldsymbol{a}\{\boldsymbol{x}(n-1)\} &\approx \boldsymbol{a}\{\hat{\boldsymbol{x}}(n-1|n-1)\} \\ &+ \left.\frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}(n-1)}\right|_{\boldsymbol{x}(n-1)=\hat{\boldsymbol{x}}(n-1|n-1)} (\boldsymbol{x}(n-1) - \hat{\boldsymbol{x}}(n-1|n-1)) \\ \boldsymbol{C}(\boldsymbol{x}(n)) &\approx \boldsymbol{C}(\hat{\boldsymbol{x}}(n|n-1)) \\ &+ \left.\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{x}(n)}\right|_{\boldsymbol{x}(n)=\hat{\boldsymbol{x}}(n|n-1)} (\boldsymbol{x}(n) - \hat{\boldsymbol{x}}(n|n-1)). \end{aligned}$$

§6.5.4 扩展卡尔曼滤波(EKF)

EKF递归估计算法：

1. 预测 $\hat{\boldsymbol{x}}(n/n-1)=\boldsymbol{a}\{(n-1)\hat{\boldsymbol{x}}(n-1/n-1)\}$
2. MMSE $\boldsymbol{P}(n/n-1)=\boldsymbol{A}(n-1)\boldsymbol{P}(n-1/n-1)\boldsymbol{A}^H(n-1)+\boldsymbol{Q}_q(n)$
3. 增益 $\boldsymbol{K}(n)=\boldsymbol{P}(n/n-1)\boldsymbol{C}^H(n)\left[\boldsymbol{C}(n)\boldsymbol{P}(n/n-1)\boldsymbol{C}^H(n)+\boldsymbol{Q}_v(n)\right]^{-1}$
4. 修正 $\hat{\boldsymbol{x}}(n/n)=\widehat{\boldsymbol{x}}(n/n-1)+\boldsymbol{K}(n)[\boldsymbol{y}(n)-\boldsymbol{C}\{\hat{\boldsymbol{x}}(n/n-1)\}]$
5. MMSE $\boldsymbol{P}(n/n)=[\boldsymbol{I}-\boldsymbol{K}(n)\boldsymbol{C}(n)]\boldsymbol{P}(n/n-1),$

$$其中:\begin{cases}\boldsymbol{A}(n-1)=\frac{\partial \boldsymbol{a}}{\partial \boldsymbol{x}(n-1)}\Big|_{\boldsymbol{x}(n-1)=\hat{\boldsymbol{x}}(n-1|n-1)}\\ \boldsymbol{C}(n)=\frac{\partial \boldsymbol{C}}{\partial \boldsymbol{x}(n)}\Big|_{\boldsymbol{x}(n)=\hat{\boldsymbol{x}}(n|n-1)}\end{cases}$$

卡尔曼滤波器：卡尔曼增益$\boldsymbol{K}(n)$，误差的协方差矩阵$\boldsymbol{P}(n/n)$，$\boldsymbol{P}(n/n-1)$与输入$\boldsymbol{x}(n)$无关，可以离线计算。

扩展卡尔曼滤波器：卡尔曼增益$\boldsymbol{K}(n)$ 需要在线计算。

另一种非线性系统中常用的卡尔曼滤波算法：**无迹卡尔曼滤波(Unscented Kalman Filtering, UKF)**，介绍从略。

总结

线性最优滤波器:

- (1)**最优FIR维纳滤波器:**用于滤波、线性预测、噪声抑制、反卷积等; 格型结构实现。
- (2)**IIR维纳滤波器:** 非因果和因果。用于滤波、线性预测、反卷积等。
- (3)**离散卡尔曼滤波器:** 对平稳和非平稳过程都适用的递归估计，可在 $n = 0$ 时刻开始工作。
- **若希望放宽对 $x(n)$ 和 $d(n)$ 的统计量的已知条件，就需自适应滤波器。**

非线性系统: 采用**扩展卡尔曼滤波（EKF）**或**无迹卡尔曼滤波（UKF）**更为有效。