



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY



Kalman滤波及拓展

数字信号处理

提纲



卡尔曼滤波



贝叶斯滤波



粒子滤波

引言

常规滤波

- 设计低通、高通、带通、带阻等滤波器，有用信号与噪声具有不同频带的条件下才能实现

维纳滤波

- 假设信号和噪声都是平稳过程的条件下，使期望输出和实际输出之间的均方误差最小来达到滤波目的

卡尔曼滤波

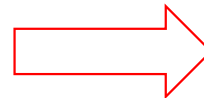
- 采用信号与噪声的**状态空间模型**，利用前一时刻的估计值和现时刻的观测值来更新对状态变量的估计，求出现时刻的估计值。
- 卡尔曼滤波不要求信号和噪声都是平稳过程的假设条件。对于每个时刻的系统扰动和观测误差（即噪声），只要对它们的统计性质作某些适当的假定，通过对含有噪声的观测信号进行处理，就能在平均的意义上，求得误差为最小的真实信号的估计值。

引言

时序AR模型转换为状态空间模型

AR(p)时间序列模型

$$x(k) = - \sum_{m=1}^p a_m x(k-m) + v(k)$$



$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{v}_{k-1}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{n}_k$$

观测值 $y(k) = x(k) + n(k)$

$$\mathbf{x}_k = \begin{bmatrix} x(k-p+1) \\ x(k-p+2) \\ \vdots \\ x(k-1) \\ x(k) \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 1 \\ -a_p & -a_{p-1} & \cdots & \cdots & -a_1 \end{bmatrix}$$

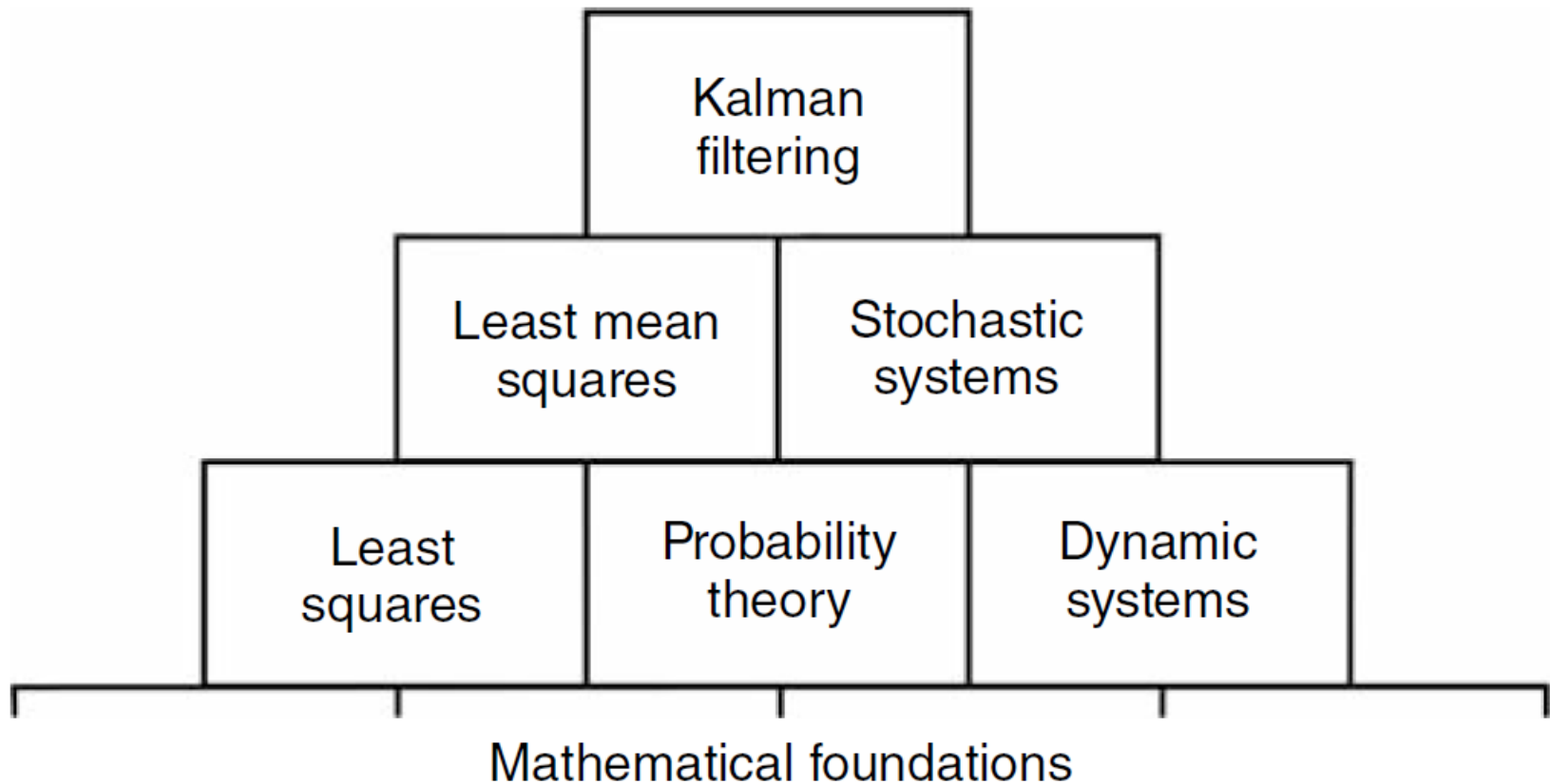
$$\mathbf{C} = [0, 0, \dots, 0, 1]$$

卡尔曼，全名Rudolf Emil Kalman，匈牙利数学家



- 1930年出生于匈牙利首都布达佩斯。1953，1954年于麻省理工学院分别获得电机工程学士及硕士学位。1957年于哥伦比亚大学获得博士学位。
- 卡尔曼滤波器，正是源于他的博士论文和1960年发表的论文《A New Approach to Linear Filtering and Prediction Problems》（线性滤波与预测问题的新方法）

卡尔曼滤波数学基础



Mohinder S. Grewal, Angus P. Andrews. Kalman filtering : theory and practice using MATLAB [Fourth edition]. Wiley, 2015



卡尔曼滤波：动态系统的状态估计

TABLE 1.1 Examples of Estimation Problems

Application	Dynamic System	Sensor Types
Process control	Chemical plant	Pressure Temperature Flow rate Gas analyzer
Flood prediction	River system	Water level Rain gauge Weather radar
Tracking	Spacecraft	Radar Imaging system
Navigation	Ship	Sextant Log Gyroscope Accelerometer GNSS ^a receiver



卡尔曼滤波的状态方程和量测方程

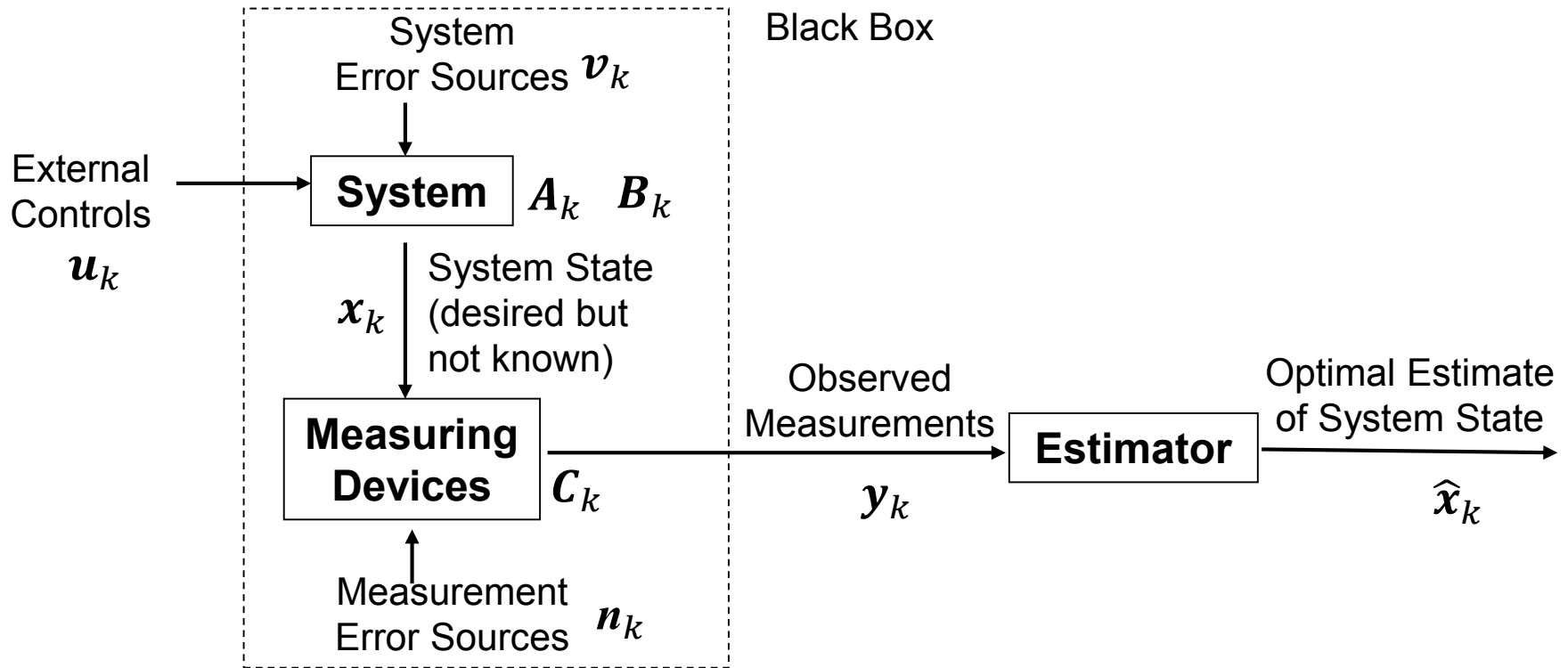
假设某系统 k 时刻的状态变量为 x_k ，状态方程和量测方程（也称为输出方程）表示为

$$x_k = A_k x_{k-1} + B_k u_{k-1} + v_{k-1}$$

$$y_k = C_k x_k + n_k$$

- A_k 为状态转移矩阵，描述系统状态由时间 $k-1$ 的状态到时间 k 的状态之间的转移；
- B_k 控制矩阵； u_k 控制向量： $B_k u_k$ 为外界激励
- C_k 为量测矩阵；
- x_k 为状态向量； y_k 为观测向量；
- v_k 为过程噪声； n_k 为量测噪声。

卡尔曼滤波的问题定义



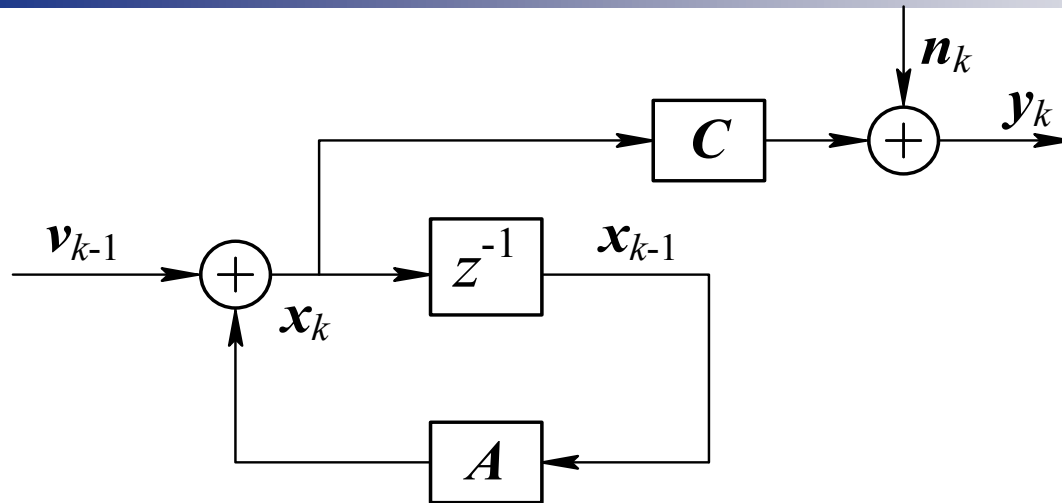
系统状态无法直接测量



需要根据测量和模型预测取得最佳估计



卡尔曼滤波的状态方程和量测方程



$$\mathbf{x}_k = A\mathbf{x}_{k-1} + \mathbf{v}_{k-1}$$

$$y_k = C\mathbf{x}_k + n_k$$

以下省略状态转移矩阵 A
和量测矩阵 C 时间下标 k

- \mathbf{v}_k 与 \mathbf{n}_k 都是零均值高斯白噪声，协方差矩阵分别是 Q 和 R

$$\mathbf{v}_k : E\{\mathbf{v}_k\} = 0, \quad E\{\mathbf{v}_k \mathbf{v}_j^T\} = Q \delta_{kj} \quad E\{\mathbf{v}_k \mathbf{n}_k^T\} = 0$$

$$\mathbf{n}_k : E\{\mathbf{n}_k\} = 0, \quad E\{\mathbf{n}_k \mathbf{n}_j^T\} = R \delta_{kj} \quad Cov\{\mathbf{x}_0, \mathbf{v}_k\} = 0; Cov\{\mathbf{x}_0, \mathbf{n}_k\} = 0$$

- 初始状态向量 \mathbf{x}_0 服从均值为 $\boldsymbol{\mu}_0$ ，方差为 $\boldsymbol{\Sigma}_0$ 的高斯分布

$$E\{(\mathbf{x}_0 - \boldsymbol{\mu}_0)(\mathbf{x}_0 - \boldsymbol{\mu}_0)^H\} = cov\{\mathbf{x}_0\} = \boldsymbol{\Sigma}_0$$



卡尔曼滤波的递推过程

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{v}_{k-1}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{n}_k$$

已知第 $k-1$ 步的状态最优估计 $\hat{\mathbf{x}}_{k-1}$ ，第 k 步测量值 \mathbf{y}_k

■ 预测过程（根据前一步的状态估计）

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A} \hat{\mathbf{x}}_{k-1} \quad \hat{\mathbf{y}}_{k|k-1} = \mathbf{C} \hat{\mathbf{x}}_{k|k-1}$$

■ 更新过程（根据当前的测量值）

$$\text{新息向量: } \boldsymbol{\alpha}_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}$$

$$\text{第}k\text{步的状态最优估计: } \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \boldsymbol{\alpha}_k$$

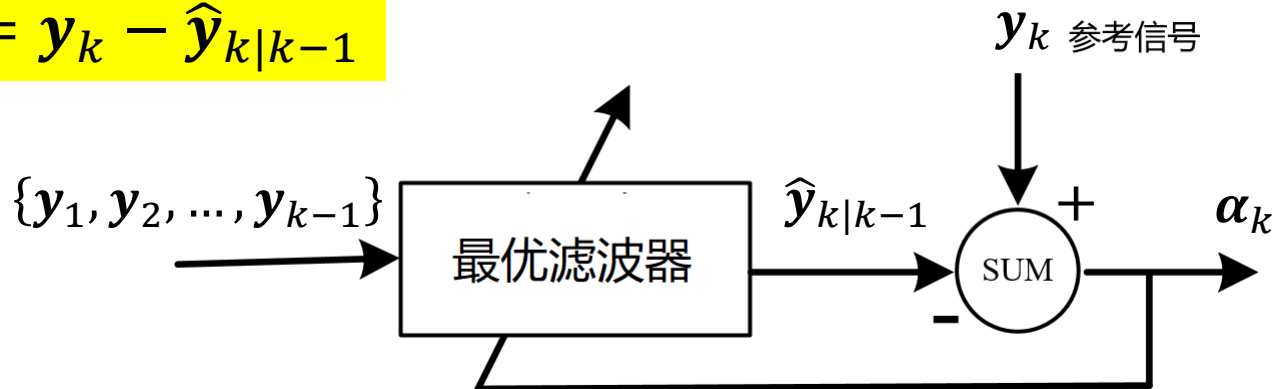
寻求卡尔曼增益 \mathbf{K}_k ,



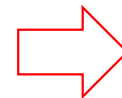
使得 $\boldsymbol{\Sigma}_k = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^H\}$ 最小化

新息向量

新息向量: $\alpha_k = \mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}$



- $\hat{\mathbf{y}}_{k|k-1}$ 是 $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{k-1}\}$ 对 \mathbf{y}_k 的最优线性预测, 则新息向量 α_k 与滤波器的输入是正交的



$$E\{\mathbf{y}_{k-j} \alpha_k^H\} = 0$$

- $\hat{\mathbf{y}}_{k|k-1}$ 是 $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{k-1}\}$ 的线性组合

$$E\{\hat{\mathbf{y}}_{k|k-1} \alpha_k^H\} = 0$$

- $\hat{\mathbf{y}}_{k|k-1} = \mathbf{C} \hat{\mathbf{x}}_{k|k-1}$

$$E\{\hat{\mathbf{x}}_{k|k-1} \alpha_k^H\} = 0$$

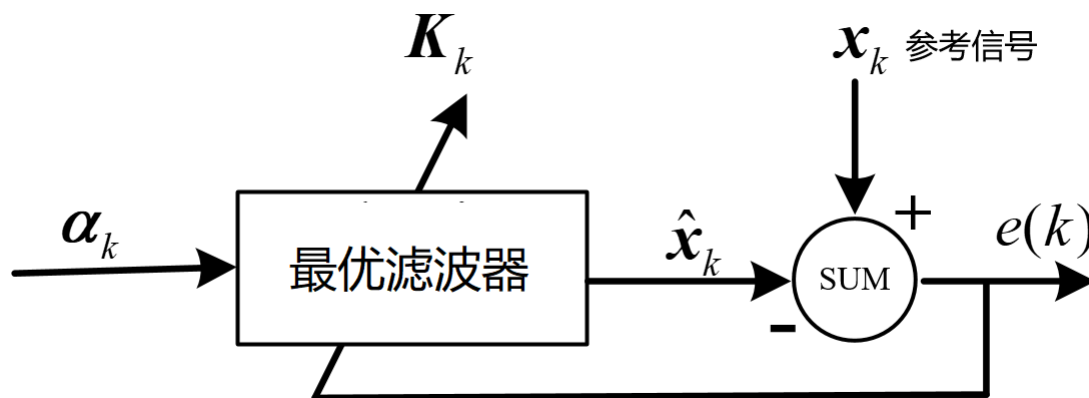
卡尔曼增益

第 k 步的状态最优估计: $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \boldsymbol{\alpha}_k$

寻求卡尔曼增益 \mathbf{K}_k ,



使得 $E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^H\}$ 最小化



滤波器输入: 新息向量 $\boldsymbol{\alpha}_k$ 的相关矩阵: $\mathbf{P}_k = E\{\boldsymbol{\alpha}_k \boldsymbol{\alpha}_k^H\}$

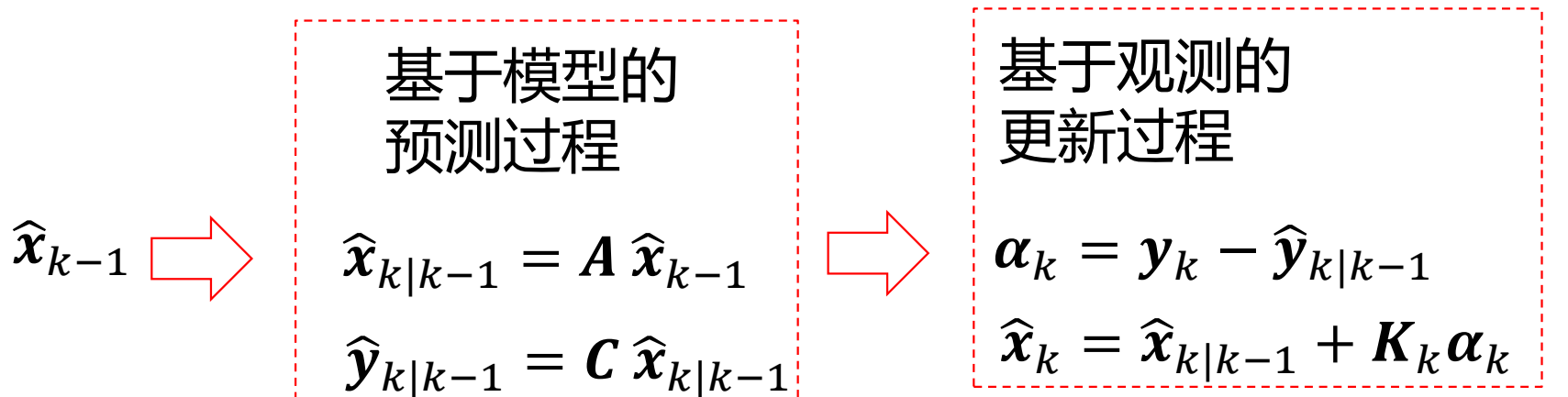
滤波器输入与参考信号的互相关矩阵: $E\{\mathbf{x}_k \boldsymbol{\alpha}_k^H\}$



$$\mathbf{K}_k = E\{\mathbf{x}_k \boldsymbol{\alpha}_k^H\} \mathbf{P}_k^{-1}$$



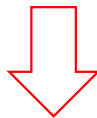
卡尔曼滤波的递推过程



$\hat{\mathbf{x}}_{k|k-1}$ ：状态的先验预测值

$\hat{\mathbf{x}}_k$ ：状态的后验估计值

定义 $\boldsymbol{\Sigma}_{k|k-1} = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^H\}$



$$\begin{aligned}\boldsymbol{\Sigma}_{k|k-1} &= E\{(\mathbf{A}\mathbf{x}_{k-1} + \mathbf{v}_{k-1} - \hat{\mathbf{x}}_{k|k-1})(\mathbf{A}\mathbf{x}_{k-1} + \mathbf{v}_{k-1} - \hat{\mathbf{x}}_{k|k-1})^H\} \\ &= \mathbf{A}\boldsymbol{\Sigma}_{k-1}\mathbf{A}^H + \mathbf{Q}\end{aligned}$$



卡尔曼滤波的递推过程

$$\alpha_k = y_k - \hat{y}_{k|k-1} = \mathbf{C}\mathbf{x}_k + \mathbf{n}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1}$$

$$\Rightarrow \mathbf{P}_k = E\{\alpha_k \alpha_k^H\} = \mathbf{C}\Sigma_{k|k-1}\mathbf{C}^H + \mathbf{R}$$

$$\because E\{\hat{\mathbf{x}}_{k|k-1} \alpha_k^H\} = 0$$

$$\Rightarrow E\{\mathbf{x}_k \alpha_k^H\} = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \alpha_k^H\} = \Sigma_{k|k-1} \mathbf{C}^H$$

$$\text{卡尔曼增益 } \mathbf{K}_k = E\{\mathbf{x}_k \alpha_k^H\} \mathbf{P}_k^{-1} = \Sigma_{k|k-1} \mathbf{C}^H (\mathbf{C}\Sigma_{k|k-1}\mathbf{C}^H + \mathbf{R})^{-1}$$

$$\text{第}k\text{步的状态最优估计: } \hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \alpha_k$$

$$\Rightarrow \Sigma_k = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^H\} = (\mathbf{I} - \mathbf{K}_k \mathbf{C}) \Sigma_{k|k-1}$$



卡尔曼滤波的递推过程

系统方程	$\begin{aligned}\mathbf{x}_k &= \mathbf{A}\mathbf{x}_{k-1} + \mathbf{v}_{k-1} \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{n}_k\end{aligned}$ <p>$\mathbf{v}_k, \mathbf{n}_k$: 零均值高斯白噪声, 协方差矩阵分别是$\mathbf{Q}$和$\mathbf{R}$</p>
初始条件	\mathbf{x}_0 : 均值为 $\boldsymbol{\mu}_0$, 方差为 $\boldsymbol{\Sigma}_0$
第 k 步 递推公式	$\begin{aligned}\hat{\mathbf{x}}_{k k-1} &= \mathbf{A}\hat{\mathbf{x}}_{k-1} && \text{状态的先验预测} \\ \boldsymbol{\alpha}_k &= \mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_{k k-1} && \text{新息向量} \\ \boldsymbol{\Sigma}_{k k-1} &= \mathbf{A}\boldsymbol{\Sigma}_{k-1}\mathbf{A}^H + \mathbf{Q} \\ \mathbf{K}_k &= \boldsymbol{\Sigma}_{k k-1}\mathbf{C}^H(\mathbf{C}\boldsymbol{\Sigma}_{k k-1}\mathbf{C}^H + \mathbf{R})^{-1} && \text{卡尔曼增益} \\ \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k k-1} + \mathbf{K}_k\boldsymbol{\alpha}_k && \text{状态的后验估计} \\ \boldsymbol{\Sigma}_k &= (\mathbf{I} - \mathbf{K}_k\mathbf{C})\boldsymbol{\Sigma}_{k k-1}\end{aligned}$

卡尔曼滤波

小结

- $\hat{x}_{k|k-1}$: $k-1$ 对 k 时刻估计值, y_k : 是 k 时刻的测量值, 估计值和测量值都存在误差, 且误差都假设满足独立的高斯分布
- Kalman滤波器: 结合了估计值和测量值得到 k 时刻更接近真值的估计结果
- Kalman滤波器引入状态空间的目的是避免了“像Wiener滤波器一样需要对过去所有 $[0, k-1]$ 时刻协方差先验知识都已知”, 而直接可以通过上一时刻即 $k-1$ 时刻的状态信息和均方误差信息就可递推得到 k 时刻的估计。
- $k-1$ 对 k 时刻的估计实际上使用到了所有前 $[0, n-1]$ 时刻的信息, 信息一直通过最小均方误差进行传递到 $k-1$ 时刻。基于此, Kalman滤波也需要先验知识, 即 -1 时刻的初始值

卡尔曼滤波的一个简单例子

房间实际温度估算

- 假定房间温度与前一时刻相同，即状态转移矩阵 $A = 1$
- 温度计直接测量房间温度，即量测矩阵 $C = 1$
- 经验预测标准差 $\sigma_v = 5$ ；温度计测量标准差 $\sigma_n = 4$

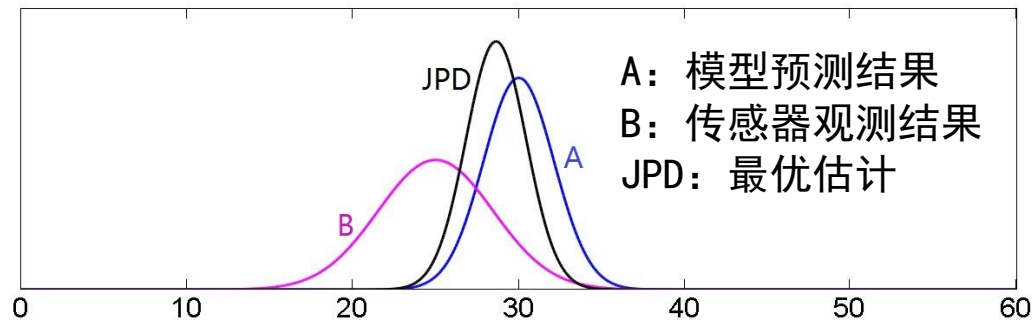
$$K_k = \Sigma_{k|k-1} C^H (C \Sigma_{k|k-1} C^H + R)^{-1} \quad \downarrow$$

$$\begin{aligned} \Sigma_{k|k-1} &= Q = \sigma_v^2; \quad K_k = \sigma_v^2 / (\sigma_v^2 + \sigma_n^2); \\ \hat{x}_k &= \hat{x}_{k-1} + K_k (y_k - \hat{x}_{k-1}); \quad \Sigma_k = (I - K_k) \sigma_v^2 \end{aligned}$$

- $K_k = \frac{25}{25+16} = 0.61$
- 设k-1时刻温度测量值23度，k时刻温度测量值25度
- 最优估计值： $\hat{x}_k = 23 + 0.61 * (25 - 23) = 24.22$ 度

卡尔曼滤波的解释

- ④ 卡尔曼滤波融合预测（状态方程）和观测（量测方程）结果，利用先验知识和新息，得到状态后验概率的最优估计



$$K_k = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_B^2}$$

$$\mu_{\text{JPD}} = \mu_A + H_k (\mu_B - \mu_A)$$

$$\sigma_{\text{JPD}}^2 = \sigma_A^2 - H_k \sigma_A^2$$

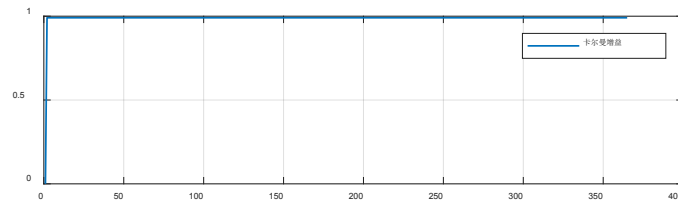
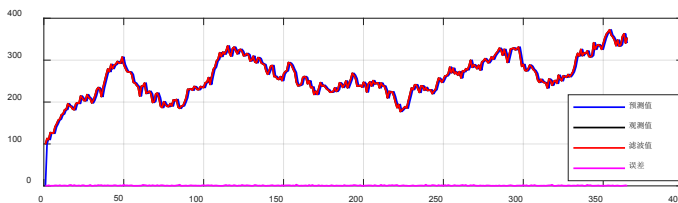
卡尔曼滤波仿真案例1



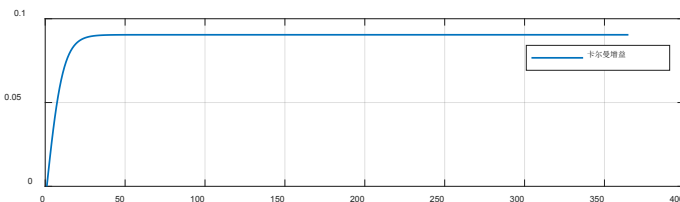
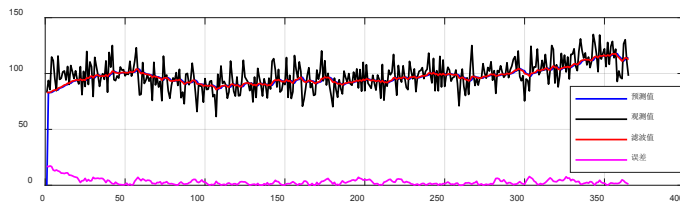
一维卡尔曼滤波器的Matlab仿真

$$\begin{cases} x(k) = x(k-1) + w(k) \\ y(k) = x(k) + v(k) \end{cases}$$

Example1.m



$W = 10 * \text{randn}(N, 1);$ % 系统过程噪声
 $V = 1 * \text{randn}(N, 1);$ % 构造观测噪声
 由于测量噪声相比于系统噪声小的多，
 所以估计值更倾向于观测值，此时的卡尔曼增益趋近于1



$W = 1 * \text{randn}(N, 1);$ % 构造过程噪声
 $V = 10 * \text{randn}(N, 1);$ % 构造观测噪声
 测量噪声相比于系统噪声要大的多，因而估计值更倾向于预测值，此时的卡尔曼增益远小于1。

卡尔曼滤波仿真案例2

2D目标跟踪

- 某质点在2维平面运动，令 z_{1t} 和 z_{2t} 表示水平和垂直位置， \dot{z}_{1t} 和 \dot{z}_{2t} 为对应速度，状态向量为 $\mathbf{Z}_t^T = (z_{1t} \ z_{2t} \ \dot{z}_{1t} \ \dot{z}_{2t})$
- 该质点在平面匀速运动，但是被随机噪声干扰；观测值为质点的平面位置，存在测量误差

$$\mathbf{Z}_t = \mathbf{A}_t \mathbf{Z}_{t-1} + \mathbf{W}_t \quad \begin{pmatrix} z_{1t} \\ z_{2t} \\ \dot{z}_{1t} \\ \dot{z}_{2t} \end{pmatrix} = \begin{Bmatrix} 1 & 0 & \Delta & 0 \\ 0 & 1 & 0 & \Delta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Bmatrix} \begin{pmatrix} z_{1,t-1} \\ z_{2,t-1} \\ \dot{z}_{1,t-1} \\ \dot{z}_{2,t-1} \end{pmatrix} + \begin{pmatrix} w_{1t} \\ w_{2t} \\ w_{3t} \\ w_{4t} \end{pmatrix}$$

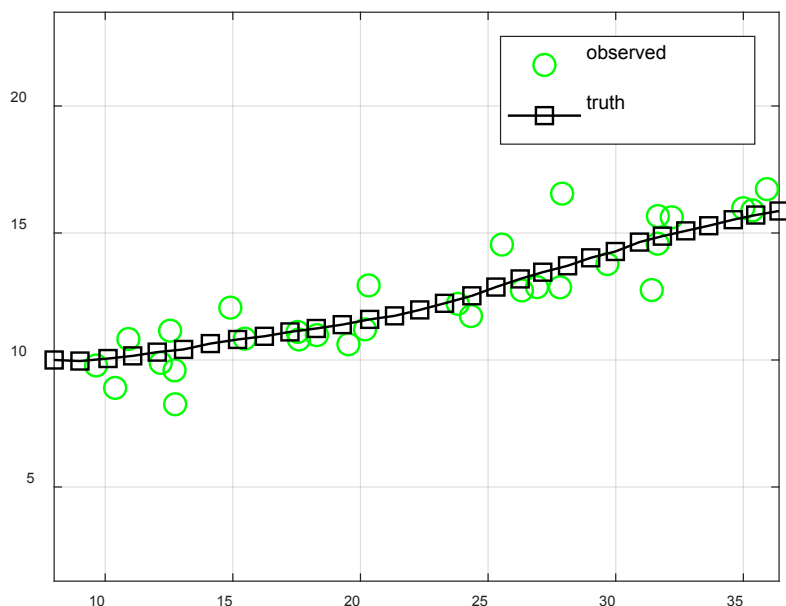
$$\mathbf{Y}_t = \mathbf{C}_t \mathbf{Z}_t + \mathbf{V}_t \quad \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{Bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{Bmatrix} \begin{pmatrix} z_{1t} \\ z_{2t} \\ \dot{z}_{1t} \\ \dot{z}_{2t} \end{pmatrix} + \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix}$$

卡尔曼滤波仿真案例2



2D目标跟踪

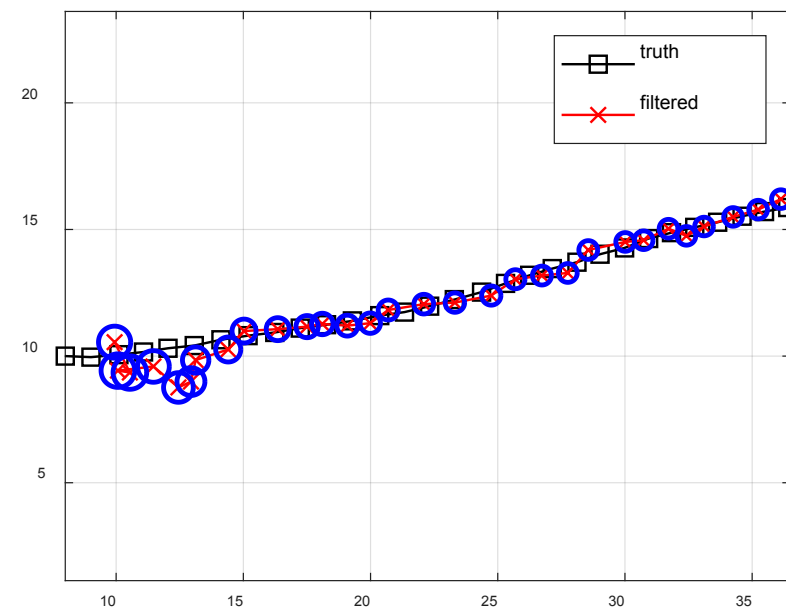
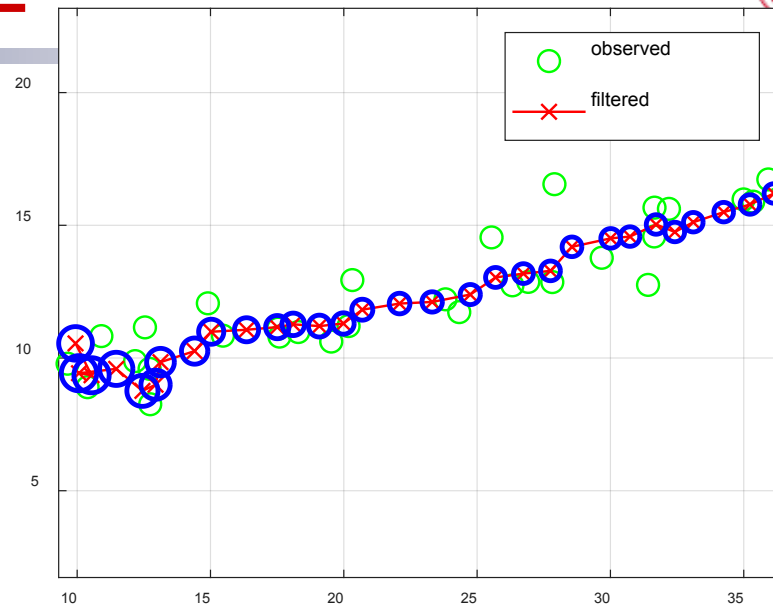
Example2.m



observed: 观测值

truth: 实际位置

filtered: 对观测值滤波结果



卡尔曼滤波图示案例

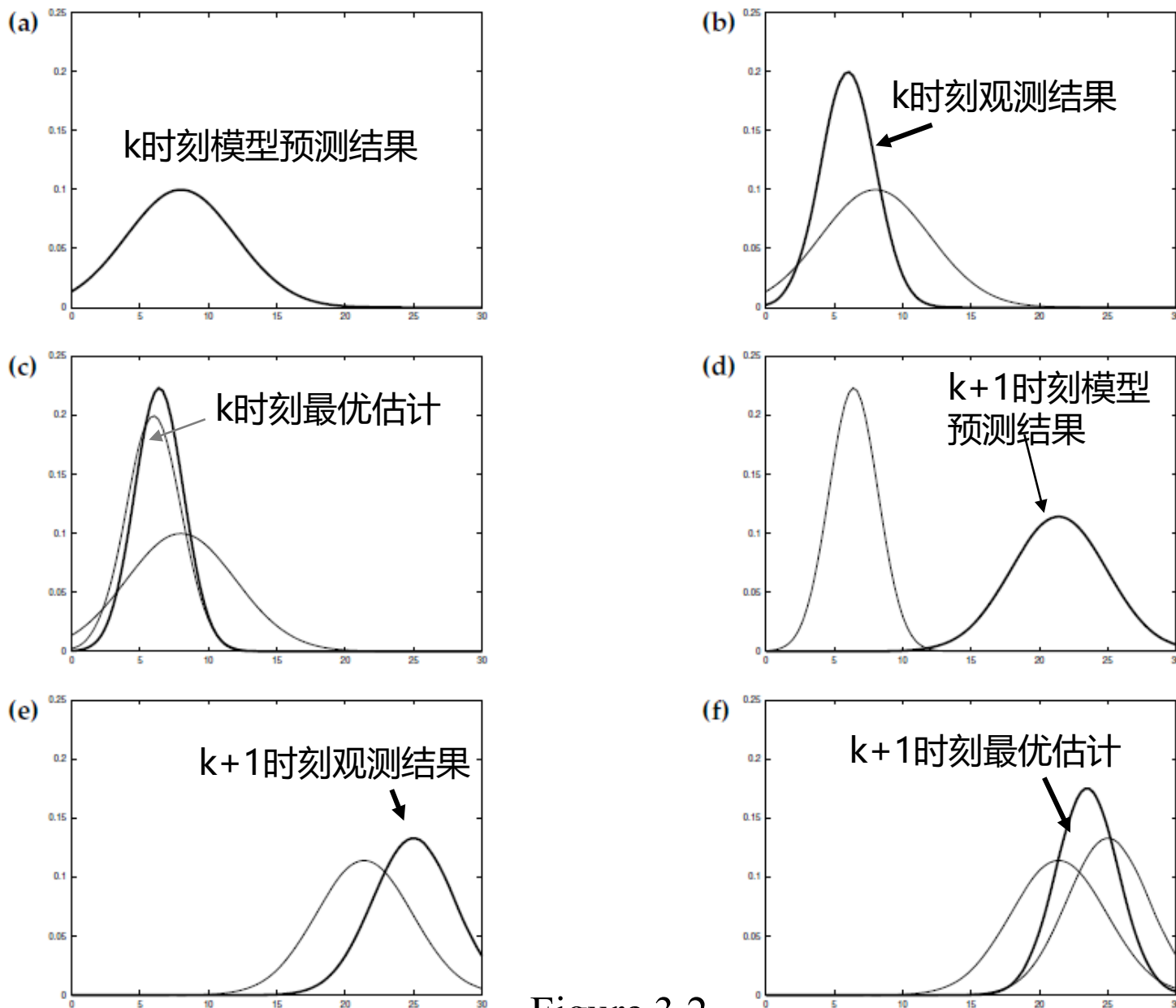


Figure 3.2

S. Thrun, W. Burgard, D. Fox. "Probabilistic Robotics[M]". 2006: The MIT Press.

扩展卡尔曼滤波

$$\mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{v}_{k-1}$$

$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{n}_k$$

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{n}_k$$

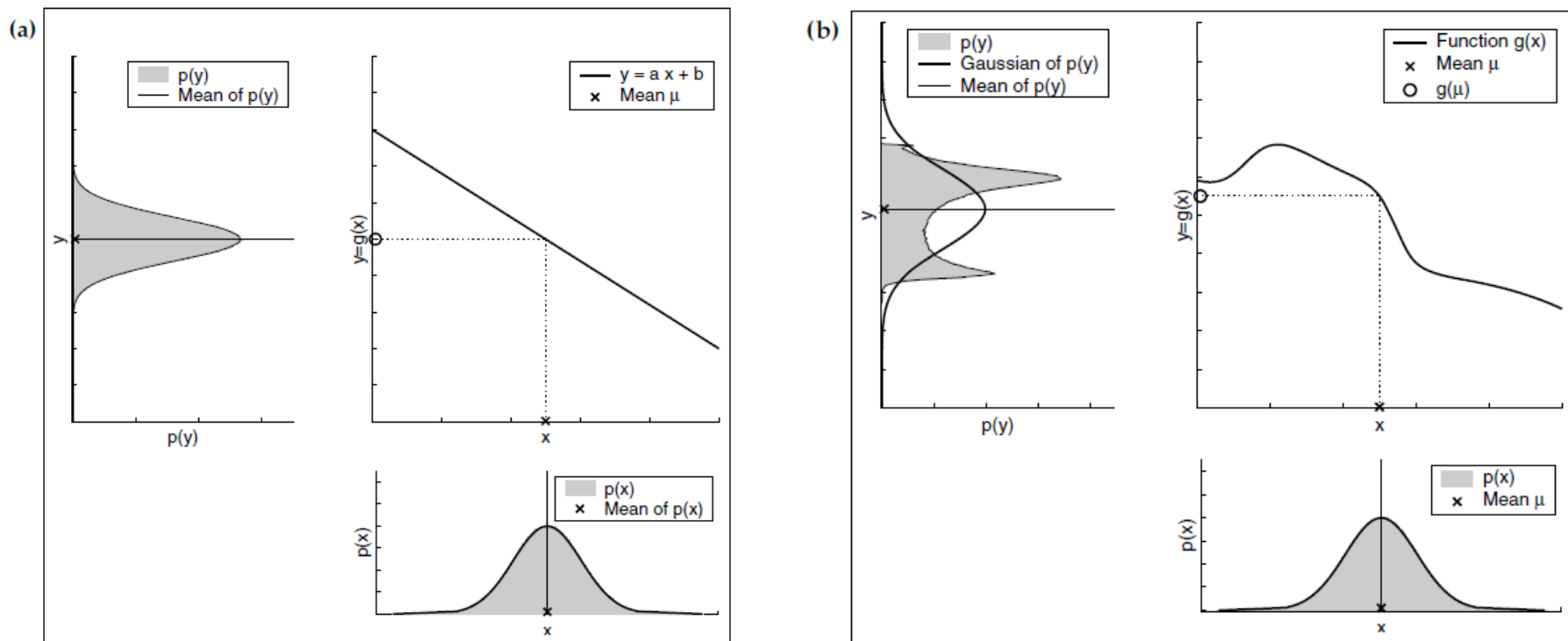


Figure 3.3 高斯随机变量通过线性与非线性系统

S. Thrun, W. Burgard, D. Fox. "Probabilistic Robotics[M]". 2006: The MIT Press.

扩展卡尔曼滤波

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1}$$

$$\mathbf{A}_k = \left. \frac{\partial \mathbf{f}_k}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_k}$$

$$\mathbf{x}_k = \mathbf{A}_k \mathbf{x}_{k-1} + \mathbf{v}_{k-1}$$

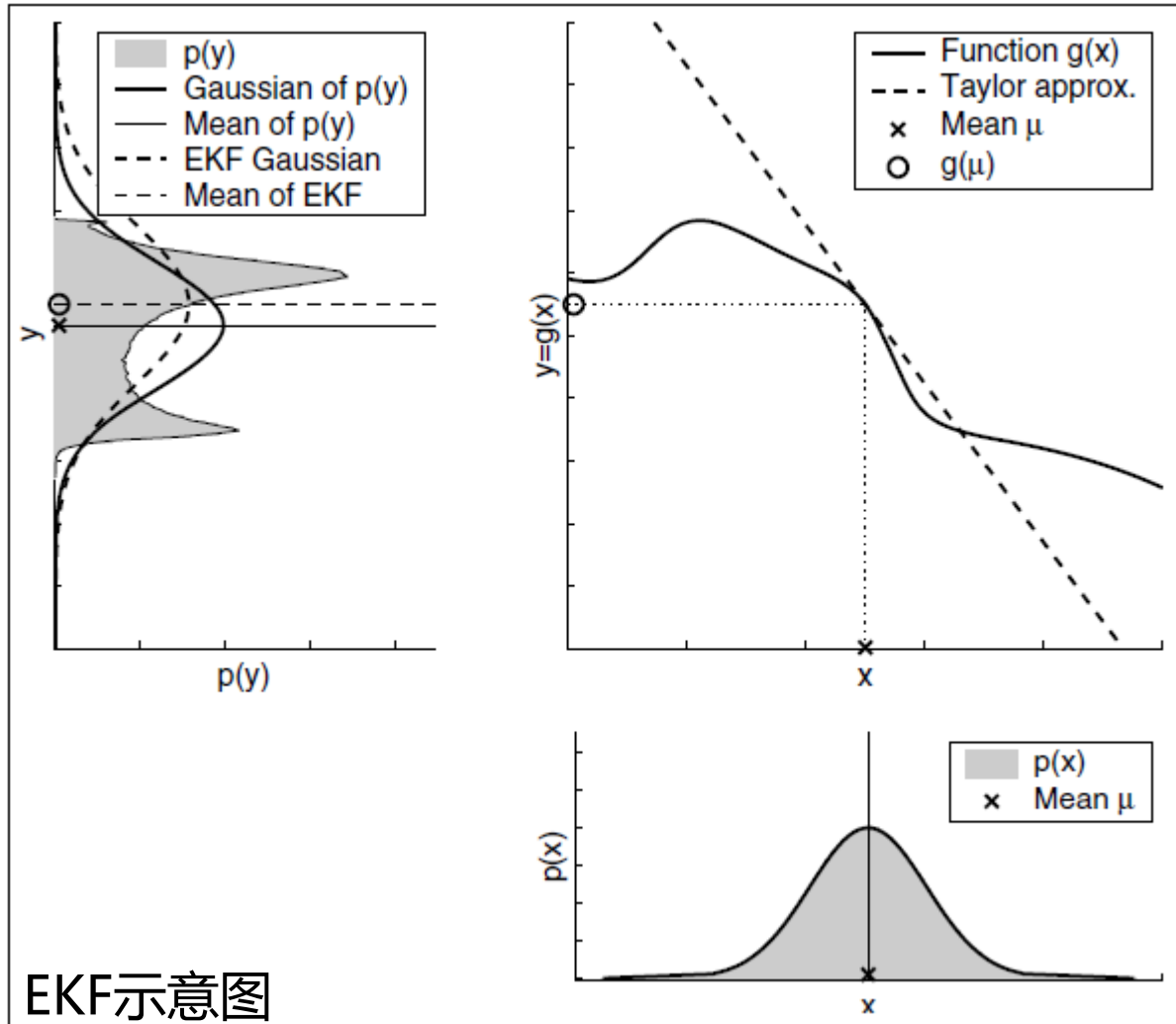


Figure 3.4 EKF示意图

扩展卡尔曼滤波

⊙ $x_k = f_k(x_{k-1})$: x_{k-1} 的方差不同带来的影响

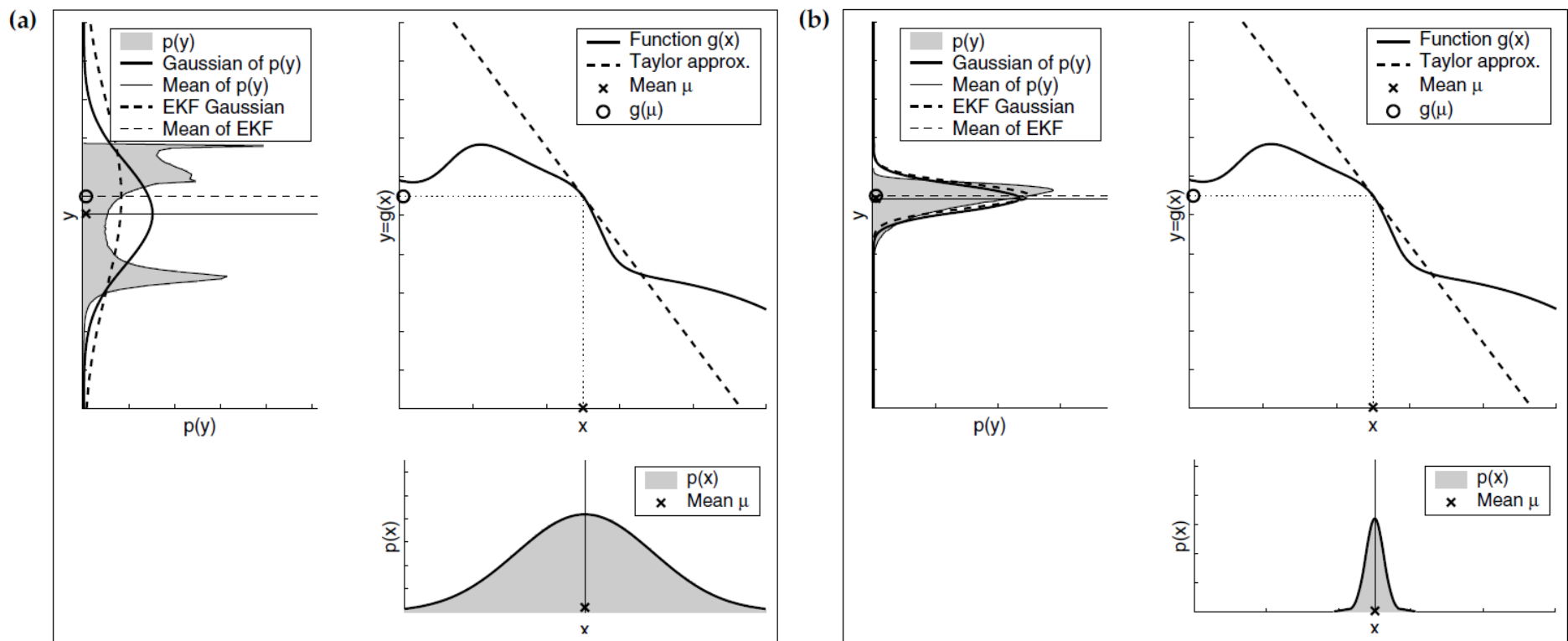



Figure 3.5

扩展卡尔曼滤波



$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}): \quad \mathbf{A}_k = \left. \frac{\partial \mathbf{f}_k}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_k}$$
 不同带来的影响

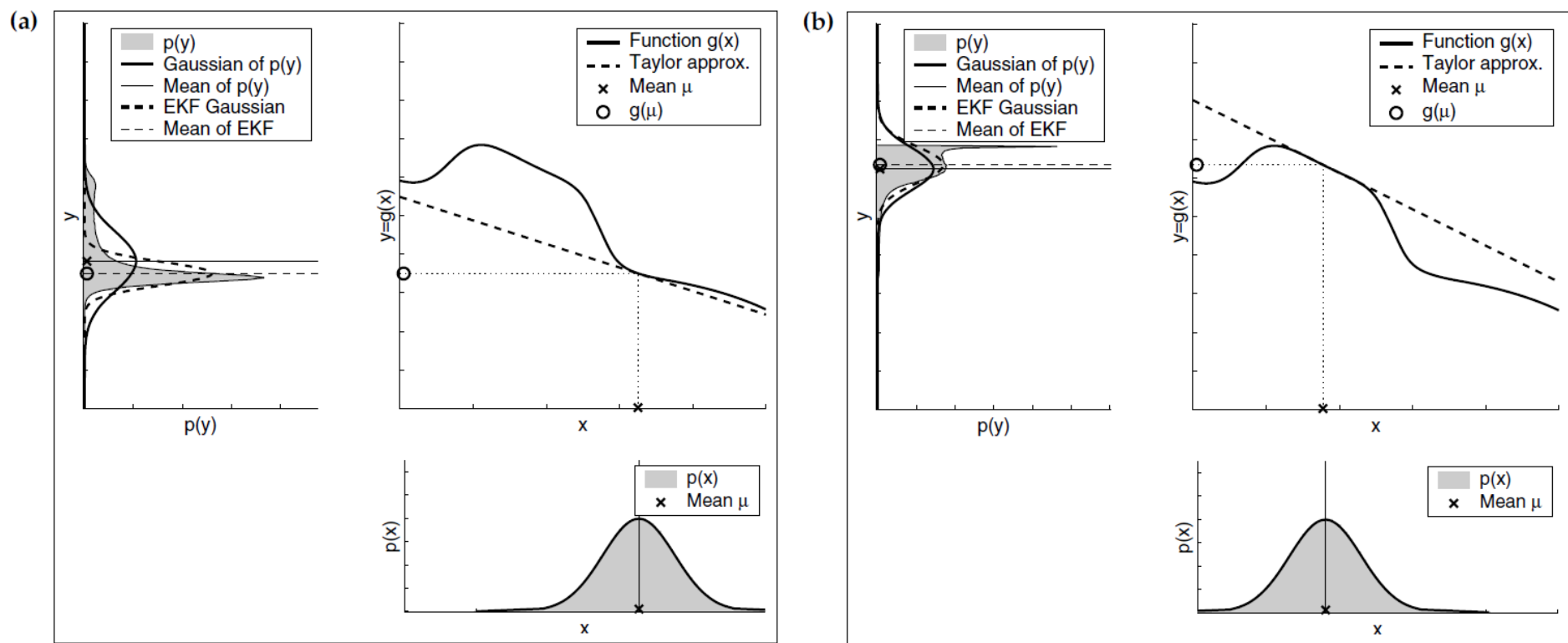


Figure 3.6

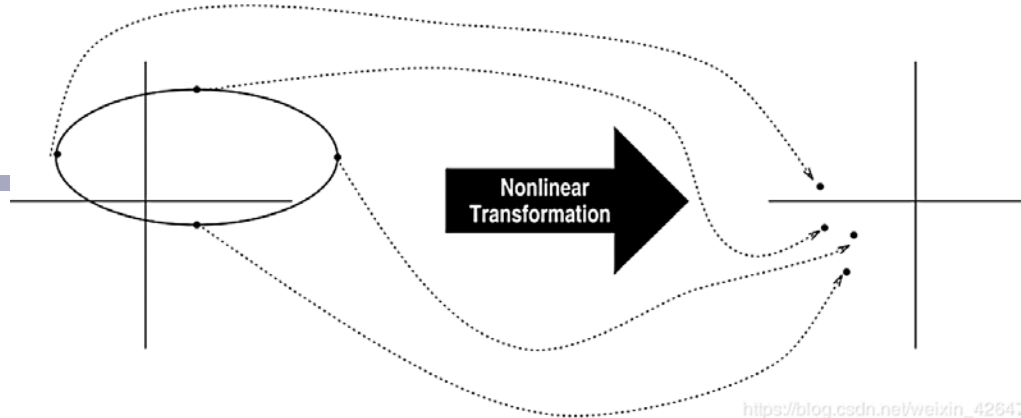


无迹卡尔曼滤波

- ❶ 无迹卡尔曼滤波（Unscented Kalman Filter, UKF），是无迹变换（Unscented Transform, UT）与标准卡尔曼滤波体系的结合，通过UT变换使非线性系统方程适用于线性假设下的标准卡尔曼体系
- ❷ 与EKF不同的是，UKF是通过无迹变换使非线性系统方程适用于线性假设下的标准Kalman滤波体系，而不是像EKF那样，必须通过线性化非线性函数实现递推滤波
- ❸ 无迹卡尔曼滤波法只需选取 $2n+1$ 个sigma点， n 为随机变量维数。相对于粒子滤波，计算效率较高

无迹卡尔曼滤波

UT变换



■ 构造Sigma点集

$$\chi_i = \begin{cases} \bar{\mathbf{x}} + (\sqrt{(n+\lambda) \cdot \mathbf{P}_x})_i, & i = 1, \dots, n \\ \bar{\mathbf{x}} - (\sqrt{(n+\lambda) \cdot \mathbf{P}_x})_i, & i = n+1, \dots, 2n \\ \bar{\mathbf{x}}, & i = 0 \end{cases}$$

n 维随机变量 x 的统计量
均值 \bar{x} 和协方差 P_x

■ 对Sigma点集做非线性变换

$$\mathbf{Y}_i = f(\chi_i), i=0,1,\dots,2n$$

■ 计算函数映射后的均值和方差

✧ 对变换后的Sigma点集做加权处理，得到输出量的均值和方差

$$\bar{\mathbf{y}} \approx \sum_{i=0}^{2n} W_i^m \mathbf{Y}_i \quad \bar{\mathbf{P}}_y \approx \sum_{i=0}^{2n} W_i^c (\mathbf{Y}_i - \bar{\mathbf{y}})(\mathbf{Y}_i - \bar{\mathbf{y}})^T$$

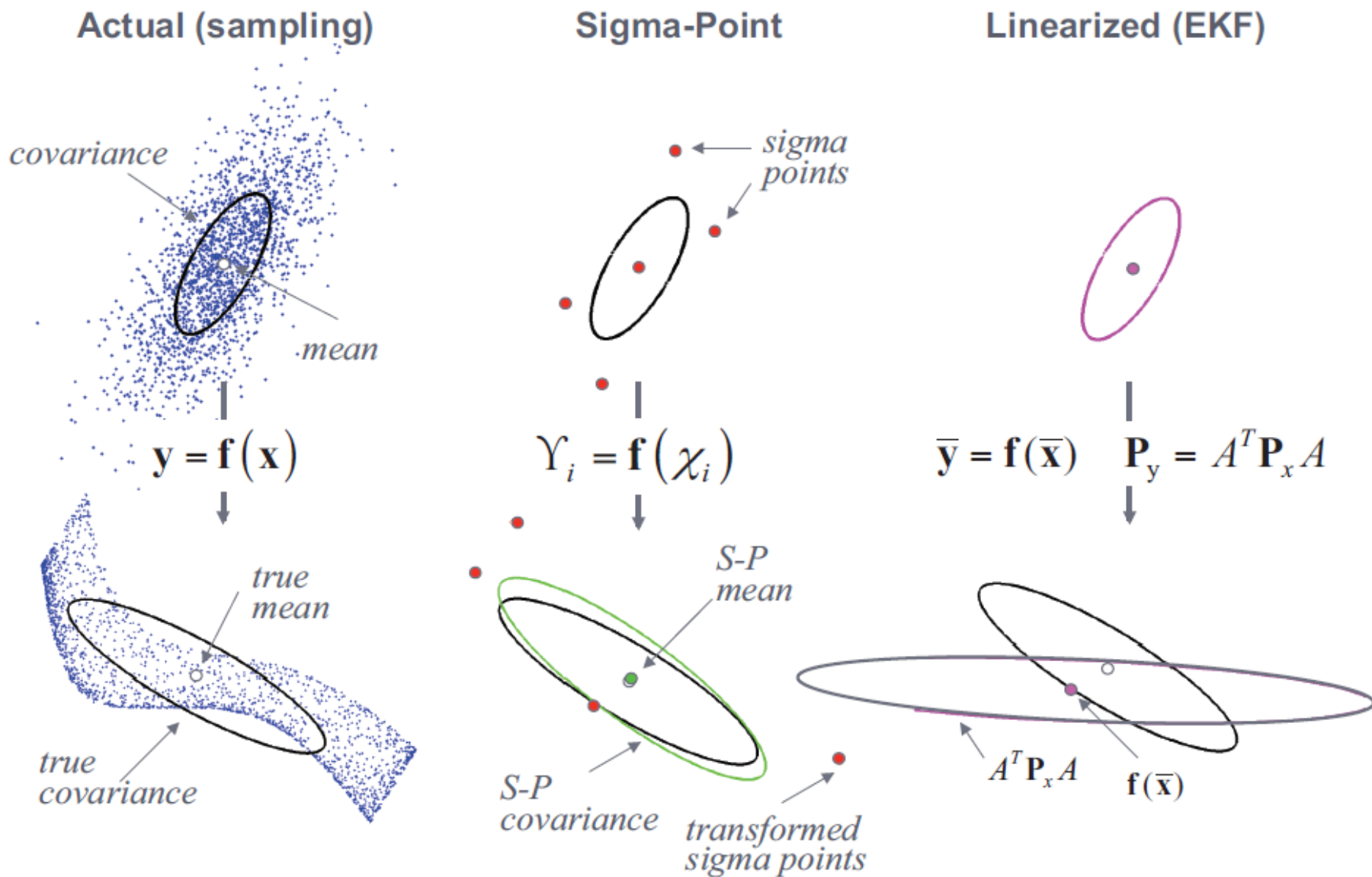


Figure 18.10 无迹变换示意

K. P. Murphy. "Machine Learning: A Probabilistic Perspective[M]". 2012: The MIT Press.

提纲



卡尔曼滤波



贝叶斯滤波



粒子滤波

贝叶斯滤波

贝叶斯公式

- 观测数据集 $\mathcal{D} = \{t_1, \dots, t_N\}$, 信号模型参数向量 \mathbf{w}
- 在观察到数据之前, 我们有一些关于模型参数 \mathbf{w} 的假设, 这以先验概率 $p(\mathbf{w})$ 的形式给出
- 在获得观测数据集 \mathcal{D} 之后, 模型参数 \mathbf{w} 的后验概率 $p(\mathbf{w}|\mathcal{D})$ 可通过如下贝叶斯公式得到

$$p(\mathbf{w}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathbf{w})p(\mathbf{w})}{p(\mathcal{D})}$$

贝叶斯滤波

贝叶斯公式(续)

- $p(\mathcal{D}|\mathbf{w})$: 似然函数(likelihood function)。它表达了在不同的参数向量下, 观测数据 \mathcal{D} 出现的可能性的
- 大小
- $p(\mathcal{D}) = \int p(\mathcal{D}|\mathbf{w})p(\mathbf{w})d\mathbf{w}$: 归一化常数, 确保后验概率 $p(\mathbf{w}|\mathcal{D})$ 积分为1

$$p(\mathbf{w}|\mathcal{D}) = \frac{p(\mathcal{D}|\mathbf{w})p(\mathbf{w})}{p(\mathcal{D})} \iff \text{posterior} \propto \text{likelihood} \times \text{prior}$$

贝叶斯滤波

研究对象

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}) \quad \text{状态方程}$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{n}_k) \quad \text{测量方程}$$

- \mathbf{x}_k 为系统状态向量, \mathbf{v}_k 为过程噪声向量, \mathbf{f}_k 为状态转移函数, 可能为非线性函数
- \mathbf{y}_k 为观测向量, \mathbf{n}_k 为测量噪声向量, \mathbf{h}_k 为测量函数, 可能为非线性函数
- 这里假设噪声向量的概率密度函数是已知的, 状态转移函数和测量函数也是已知的

贝叶斯滤波

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1})$$

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{n}_k)$$

④ 贝叶斯滤波：根据已测数据 $\mathbf{y}_{1:k} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k\}$ ，
计算当前状态 \mathbf{x}_k 的最优估计

■ 若考虑均方误差准则， \mathbf{x}_k 的 **后验条件期望** 是最优估计

④ 贝叶斯滤波通过递推求解 \mathbf{x}_k 的 **后验概率** $p(\mathbf{x}_k | \mathbf{y}_{1:k})$

■ 假设 $k-1$ 时刻的状态后验概率密度 $p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1})$ 已知

■ 递推过程包含 **预测** 和 **更新** 两个过程

贝叶斯滤波

预测过程

- 利用系统状态转移方程预测状态的先验概率密度

$$p(\mathbf{x}_k | \mathbf{y}_{1:k-1})$$

$$p(\mathbf{x}_k | \mathbf{y}_{1:k-1}) = \int p(\mathbf{x}_k, \mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \quad \text{概率加和规则}$$

$$= \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{1:k-1}) p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1} \quad \text{概率乘积规则}$$

$$= \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{x}_{k-1}$$



$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1})$$

系统状态转移一阶马尔科夫过程的假设，即 k 时刻状态 \mathbf{x}_k 仅由 $k-1$ 时刻的状态 \mathbf{x}_{k-1} 决定

贝叶斯滤波



更新过程

$$\mathbf{y}_k = \mathbf{h}_k(\mathbf{x}_k, \mathbf{n}_k)$$

- 利用最新的测量值 \mathbf{y}_k ，对先验概率密度 $p(\mathbf{x}_k|\mathbf{y}_{1:k-1})$ 进行修正，得到后验概率密度 $p(\mathbf{x}_k|\mathbf{y}_{1:k})$

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{y}_{1:k}) &= \frac{p(\mathbf{y}_k|\mathbf{x}_k, \mathbf{y}_{1:k-1})p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1})} \\ &= \frac{p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})}{p(\mathbf{y}_k|\mathbf{y}_{1:k-1})} \end{aligned}$$

贝叶斯公式

- $p(\mathbf{y}_k|\mathbf{x}_k)$: 似然函数
- $p(\mathbf{y}_k|\mathbf{y}_{1:k-1}) = \int p(\mathbf{y}_k|\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{1:k-1})d\mathbf{x}_k$: 归一化常数

贝叶斯滤波

- ❶ 贝叶斯滤波通过预测过程和更新过程迭代进行，从而实现最优状态估计
- ❷ 贝叶斯滤波是最优滤波器的系统性框架，可应用于高斯和非高斯分布的一般情况，但是，贝叶斯滤波一般只能采用数值逼近算法实现
- ❸ 当系统状态方程和量测方程为线性方程，噪声变量满足高斯分布，这种情况贝叶斯滤波的最优解是卡尔曼滤波

贝叶斯滤波

高斯分布的一些性质

- 一元高斯随机变量 x , 均值为 μ , 方差为 σ^2

$$\mathcal{N}(x|\mu, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma^2}(x - \mu)^2\right\}$$

- N 维向量 \mathbf{x} , 多元高斯分布

$$\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{\frac{N}{2}}} \frac{1}{|\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T (\boldsymbol{\Sigma})^{-1} (\mathbf{x} - \boldsymbol{\mu})\right\}$$

$\boldsymbol{\mu}$: N 维均值向量

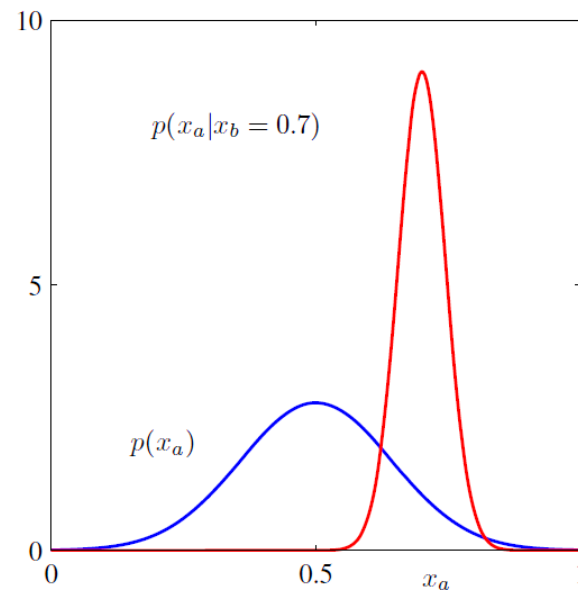
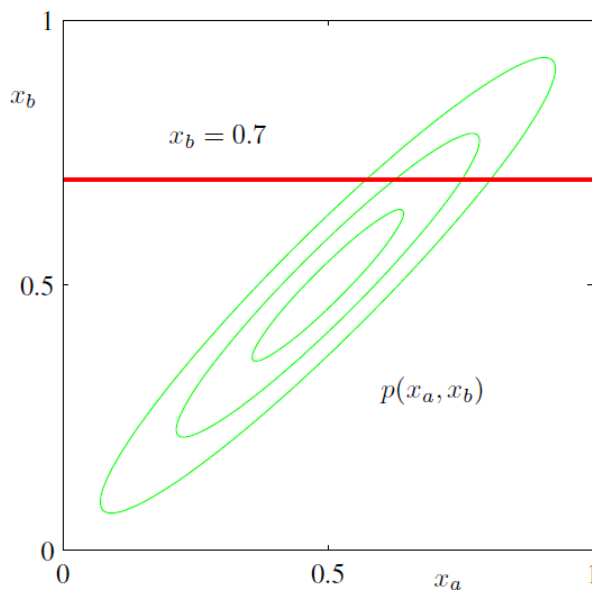
$\boldsymbol{\Sigma}$: $N \times N$ 的协方差矩阵

$|\boldsymbol{\Sigma}|$: $\boldsymbol{\Sigma}$ 的行列式

贝叶斯滤波

高斯分布的一些性质

- 如果联合分布 $p(\mathbf{x}_a, \mathbf{x}_b)$ 是高斯分布, 则条件概率分布 $p(\mathbf{x}_a|\mathbf{x}_b)$ 也是高斯分布
- 边缘概率分布 $p(\mathbf{x}_a) = \int p(\mathbf{x}_a, \mathbf{x}_b) d\mathbf{x}_b$ 也是高斯分布
- 条件概率分布 $p(\mathbf{x}_a|\mathbf{x}_b)$ 的均值是 \mathbf{x}_b 的线性函数, 而协方差与 \mathbf{x}_b 无关



贝叶斯滤波

高斯分布的一些性质

- 给定一个高斯边缘分布 $p(\mathbf{x})$ 和一个高斯条件分布 $p(\mathbf{y}|\mathbf{x})$, 其中 $p(\mathbf{y}|\mathbf{x})$ 的均值是 \mathbf{x} 的线性函数, 协方差与 \mathbf{x} 无关

$$p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1}) \quad p(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{y}|\mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{L}^{-1})$$

※ $\boldsymbol{\Lambda}$ 和 \mathbf{L} 是精度矩阵, 协方差矩阵的逆矩阵

- 边缘概率分布 $p(\mathbf{y})$ 均值向量和协方差矩阵

$$\mathbb{E}[\mathbf{y}] = \mathbf{A}\boldsymbol{\mu} + \mathbf{b} \quad \text{cov}[\mathbf{y}] = \mathbf{L}^{-1} + \mathbf{A}\boldsymbol{\Lambda}^{-1}\mathbf{A}^T$$

- 条件概率分布 $p(\mathbf{x}|\mathbf{y})$ 的均值向量和协方差矩阵

$$\begin{aligned} \mathbb{E}[\mathbf{x}|\mathbf{y}] &= (\boldsymbol{\Lambda} + \mathbf{A}^T\mathbf{L}\mathbf{A})^{-1}\{\mathbf{A}^T\mathbf{L}(\mathbf{y} - \mathbf{b}) + \boldsymbol{\Lambda}\boldsymbol{\mu}\} \\ \text{cov}[\mathbf{x}|\mathbf{y}] &= (\boldsymbol{\Lambda} + \mathbf{A}^T\mathbf{L}\mathbf{A})^{-1} \end{aligned}$$

贝叶斯滤波

卡尔曼滤波公式的导出

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{v}_{k-1}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{n}_k$$

\mathbf{v}_k 和 \mathbf{n}_k 为零均值高斯白噪声,
其协方差矩阵分别为 \mathbf{Q} 和 \mathbf{R} :

$$p(\mathbf{v}_k) = \mathcal{N}(\mathbf{v}_k|0, \mathbf{Q}),$$

$$p(\mathbf{n}_k) = \mathcal{N}(\mathbf{n}_k|0, \mathbf{R})$$

■ 预测过程

 $k-1$ 时刻的状态后验概率密度

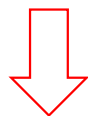
$$p(\mathbf{x}_{k-1}|\mathbf{y}_{1:k-1}) = \mathcal{N}(\mathbf{x}_{k-1}|\boldsymbol{\mu}_{k-1}, \boldsymbol{\Sigma}_{k-1}) \quad \text{已知}$$

$$p(\mathbf{x})$$

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k|\mathbf{A}\mathbf{x}_{k-1}, \mathbf{Q})$$

已知

$$p(\mathbf{y}|\mathbf{x})$$



$$p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \mathcal{N}(\mathbf{x}_k|\mathbf{A}\boldsymbol{\mu}_{k-1}, \boldsymbol{\Sigma}_{k|k-1})$$

$$\boldsymbol{\Sigma}_{k|k-1} = \mathbf{Q} + \mathbf{A}\boldsymbol{\Sigma}_{k-1}\mathbf{A}^T$$

$$p(\mathbf{y})$$

贝叶斯滤波

卡尔曼滤波公式的导出

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{v}_{k-1}$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{n}_k$$



\mathbf{v}_k 和 \mathbf{n}_k 为零均值高斯白噪声,
其协方差矩阵分别为 \mathbf{Q} 和 \mathbf{R} :

$$p(\mathbf{v}_k) = \mathcal{N}(\mathbf{v}_k|0, \mathbf{Q}),$$

$$p(\mathbf{n}_k) = \mathcal{N}(\mathbf{n}_k|0, \mathbf{R})$$

■ 更新过程

$$p(\mathbf{x}_k|\mathbf{y}_{1:k-1}) = \mathcal{N}(\mathbf{x}_k|\mathbf{A}\boldsymbol{\mu}_{k-1}, \boldsymbol{\Sigma}_{k|k-1})$$

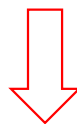
已知

$p(\mathbf{x})$

已知

$p(\mathbf{y}|\mathbf{x})$

$$p(\mathbf{y}_k|\mathbf{x}_k) = \mathcal{N}(\mathbf{y}_k|\mathbf{C}\mathbf{x}_k, \mathbf{R})$$



$$\boldsymbol{\Sigma}_k = (\boldsymbol{\Sigma}_{k|k-1}^{-1} + \mathbf{C}^T \mathbf{R}^{-1} \mathbf{C})^{-1}$$

$p(\mathbf{x}|\mathbf{y})$

$$\boldsymbol{\mu}_k = \boldsymbol{\Sigma}_k \{ \mathbf{C}^T \mathbf{R}^{-1} \mathbf{y}_k + \boldsymbol{\Sigma}_{k|k-1}^{-1} \mathbf{A} \boldsymbol{\mu}_{k-1} \}$$

k 时刻的后验高斯概率密度 $p(\mathbf{x}_k|\mathbf{y}_{1:k}) = \mathcal{N}(\mathbf{x}_k|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$

提纲



卡尔曼滤波



贝叶斯滤波



粒子滤波

粒子滤波

- ④ 当状态方程和测量方程为线性方程，噪声变量满足高斯分布，贝叶斯滤波的最优解就是卡尔曼滤波。但对于一般的非线性，非高斯系统，很难得到后验概率的解析解。粒子滤波(Particle Filter PF)是通过数值逼近实现贝叶斯滤波的一种途径
- ④ 粒子滤波的思想是通过 x_k 的一组样本(称为粒子)来逼近后验概率密度，其建立在蒙特卡洛模拟和序列重要性采样 (Sequential Importance Sampling-SIS) 方法基础上进行

粒子滤波

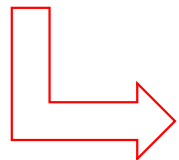
蒙特卡洛模拟与重要性采样

- 后验概率密度 $p(\mathbf{x}_k | \mathbf{y}_{1:k})$ ，可以通过蒙特卡洛模拟获得：

$$\hat{p}(\mathbf{x}_k | \mathbf{y}_{1:k}) = \frac{1}{N_s} \sum_{i=1}^{N_s} \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}) \approx p(\mathbf{x}_k | \mathbf{y}_{1:k})$$

- 问题是后验概率 $p(\mathbf{x}_k | \mathbf{y}_{1:k})$ 未知，无法采样，因此，通过选用一个已知PDF $q(\mathbf{x})$ 产生样本，这个选用的PDF称为**重要性概率密度**，从中产生的样本称为重要性采样

$$\{\mathbf{x}^{(i)} \sim q(\mathbf{x}), \quad i = 1, 2, \dots, N_s\}$$



$$w^{(i)} \propto \frac{p(\mathbf{x}^{(i)})}{q(\mathbf{x}^{(i)})}; \quad \sum_{i=1}^{N_s} w^{(i)} = 1 \quad \Rightarrow \quad p(\mathbf{x}) \approx \sum_{i=1}^{N_s} w^{(i)} \delta(\mathbf{x} - \mathbf{x}^{(i)})$$

$w^{(i)}$ 为对应第 i 个粒子的归一化的权重

粒子滤波

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}) \\ \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k, \mathbf{n}_k) \end{aligned}$$

序列重要性采样

- 递推形式的重要性采样方法
- 时刻 k 之前的粒子逼近已知: $\mathbf{x}_{0:k-1}^{(i)} \sim q(\mathbf{x}_{0:k-1} | \mathbf{y}_{1:k-1})$
- 噪声概率密度函数已知: $\rightarrow p(\mathbf{x}_k | \mathbf{x}_{k-1})$ 和 $p(\mathbf{y}_k | \mathbf{x}_k)$ 已知
- 粒子权重的递推关系

$$w_k^{(i)} \propto \frac{p(\mathbf{x}_{0:k}^{(i)} | \mathbf{y}_{1:k})}{q(\mathbf{x}_{0:k}^{(i)} | \mathbf{y}_{1:k})} \propto \frac{p(\mathbf{y}_k | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}) p(\mathbf{x}_{0:k-1}^{(i)} | \mathbf{y}_{1:k-1})}{q(\mathbf{x}_k^{(i)} | \mathbf{x}_{0:k-1}^{(i)}, \mathbf{y}_{1:k}) q(\mathbf{x}_{0:k-1}^{(i)} | \mathbf{y}_{1:k-1})}$$

$$\Rightarrow w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{y}_k | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k)}$$

粒子滤波

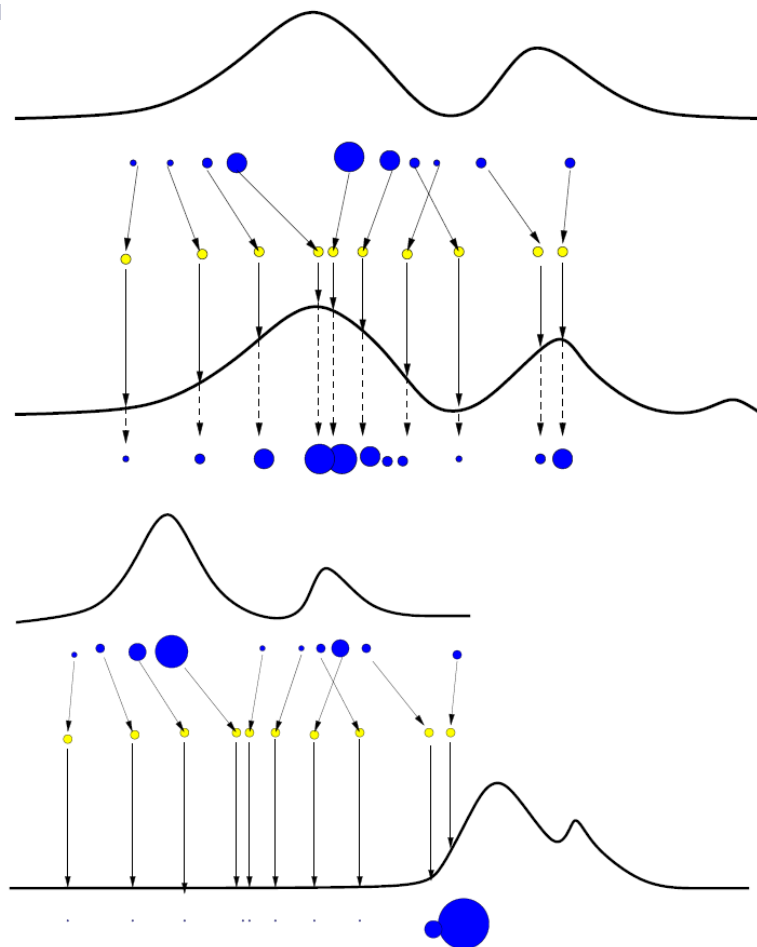
基于序列重要性采样的粒子滤波算法

初始条件	<p>给定$k = 0$时刻初始分布$q(x_0 y_0, x_{-1}) = p(x_0)$，产生$N_s$个样本$x_0^{(i)}$，权值$w_0^{(i)} = 1/N_s$</p>
粒子更新和滤波公式	<p>k时刻（从1开始），获得观测值y_k</p> <p>对于$i = 1, 2, \dots, N_s$</p> <p>产生新样本：$x_k^{(i)} \sim q(x_k x_{k-1}^{(i)}, y_k)$</p> <p>计算新的权值$w_k^{(i)}$ $\leftarrow w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(y_k x_k^{(i)})p(x_k^{(i)} x_{k-1}^{(i)})}{q(x_k^{(i)} x_{k-1}^{(i)}, y_k)}$</p> <p>权值归一化,逼近后验概率</p> $\hat{p}(x_k y_{1:k}) = \sum_{i=1}^{N_s} w_k^{(i)} \delta(x_k - x_k^{(i)})$ <p>均方误差最小化的粒子滤波输出</p> $\hat{x}_{k k} = \sum_{i=1}^{N_s} w_k^{(i)} x_k^{(i)}$ <p>$k \leftarrow k + 1$,得到新观测值y_k，进入下一次循环</p>

粒子滤波

粒子权重退化问题

- 经过几次迭代以后，很多粒子的权重都变得很小，可以忽略了，只有少数粒子的权重比较大
- 大量的计算浪费在对估计后验滤波概率分布几乎不起作用的粒子上，使得估计性能下降



➡ **重采样**，权重大的粒子复制的次数多

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(y_k | x_k^{(i)}) p(x_k^{(i)} | x_{k-1}^{(i)})}{q(x_k^{(i)} | x_{k-1}^{(i)}, y_k)}$$

粒子滤波

重采样过程

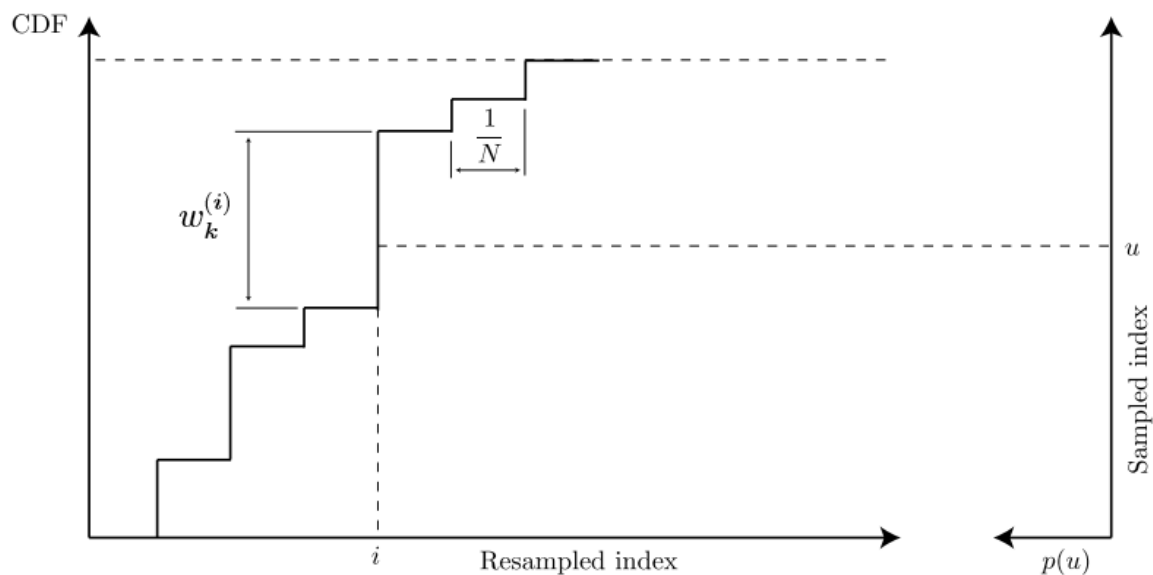
(1) 对归一化的权值 $w_k^{(i)}$ 计算概率累计和 (Matlab 中为 cumsum 函数) $CDF(i)$:

$$CDF(i) = \sum_{m=1}^i w_k^{(m)}, \quad i = 1, 2, \dots, N_s$$

(2) 取初始值 $i = 1$, 随机数 $u_1 = rand[0, 1/N_s]$

(3) 重采样循环:

```
for j = 1: N_s
    u = u_1 + (j - 1)/N_s
    while u > CDF(i)
        i = i + 1;
    end
    x_k^{(j)} = x_k^{(i)}
end
```



重采样以后, 粒子集 $x_k^{(j)}$ 所有的粒子权重一样, 都是 $1/N_s$

粒子滤波

重采样过程

$$p(\mathbf{x}_k | \mathbf{y}_{1:k}) = \sum_{i=1}^{N_s} w_k^{(i)} \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)})$$

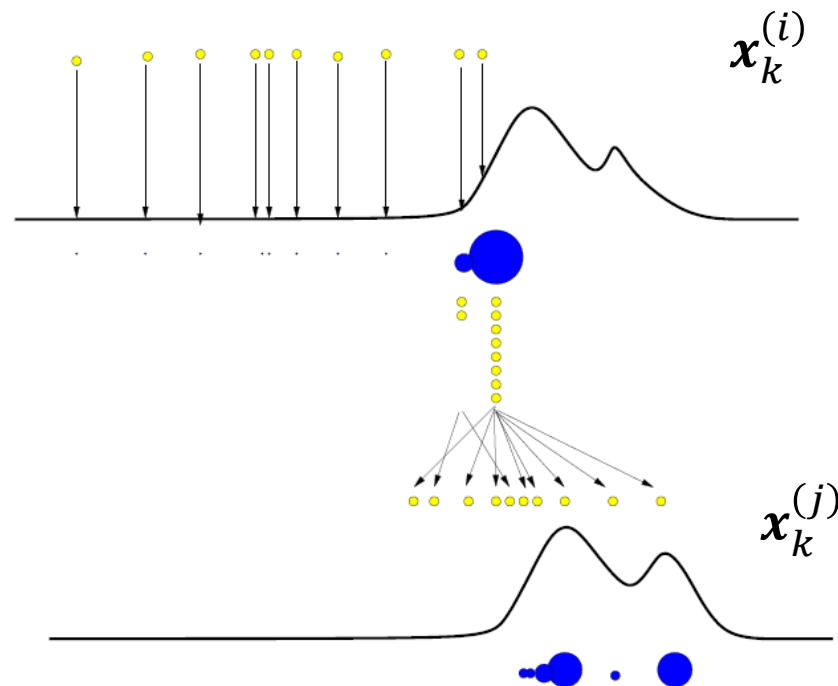
$\mathbf{x}_k^{(i)}$: k 时刻经过序列重要性采样递推得到的粒子

$\mathbf{x}_k^{(j)}$: k 时刻重采样以后的粒子

$$\tilde{p}(\mathbf{x}_k | \mathbf{y}_{1:k}) = \sum_{i=1}^{N_s} \frac{1}{N_s} \delta(\mathbf{x}_k - \mathbf{x}_k^{(j)}) = \sum_{i=1}^{N_s} \frac{n_i}{N_s} \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)})$$

n_i 是权值较大粒子 $\mathbf{x}_k^{(i)}$ 在产生新的粒子集 $\mathbf{x}_k^{(j)}$ 时被复制的次数

粒子集 $\mathbf{x}_k^{(j)}$ 所有的粒子权重一样，都是 $1/N_s$ ，只是有的粒子多出现了 n_i 次



粒子滤波

重要性概率密度函数的选择

- 在时刻 k , 用重要性概率密度函数 $q(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k)$ 产生新的粒子集 $\mathbf{x}_k^{(i)}$: $\mathbf{x}_k^{(i)} \sim q(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k)$
- 可令: $q(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k) = p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(i)})$

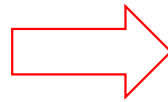
$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p(\mathbf{y}_k | \mathbf{x}_k^{(i)}) p(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})}{q(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, \mathbf{y}_k)} \quad \Rightarrow \quad w_k^{(i)} = w_{k-1}^{(i)} p(\mathbf{y}_k | \mathbf{x}_k^{(i)})$$

- 采用如上权系数更新公式的粒子滤波算法也称为SIR粒子滤波(Sampling Importance Resampling)

粒子滤波

SIR粒子滤波

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}) \\ \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k, \mathbf{n}_k) \end{aligned}$$



$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1} \\ \mathbf{y}_k &= \mathbf{h}_k(\mathbf{x}_k) + \mathbf{n}_k \end{aligned}$$

■ 非线性系统方程简化为加性噪声模型

■ $w_k^{(i)} = w_{k-1}^{(i)} p(\mathbf{y}_k | \mathbf{x}_k^{(i)})$ 。当 \mathbf{n}_k 为零均值的高斯白噪声，协方差矩阵权 Σ_n ，似然函数 $p(\mathbf{y}_k | \mathbf{x}_k^{(i)})$ 可写为多元高斯密度函数形式

$$p(\mathbf{y}_k | \mathbf{x}_k^{(i)}) = \mathcal{N}(\mathbf{y}_k | \mathbf{h}_k(\mathbf{x}_k^{(i)}), \Sigma_n)$$

$$= \frac{1}{(2\pi)^{\frac{N}{2}}} \frac{1}{|\Sigma_n|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \left(\mathbf{y}_k - \mathbf{h}_k(\mathbf{x}_k^{(i)}) \right)^T (\Sigma_n)^{-1} \left(\mathbf{y}_k - \mathbf{h}_k(\mathbf{x}_k^{(i)}) \right) \right\}$$



上海交通大学
SHANGHAI JIAO TONG UNIVERSITY



谢谢！

