- **卡尔曼滤波**
  - 多维卡尔曼滤波

➢ $\boldsymbol{Q}_k \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{Q}_k})$，$\boldsymbol{R}_k \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\boldsymbol{R}_k})$，$\boldsymbol{x}_{k-1}^{+} \sim \mathcal{N}(\boldsymbol{\mu}_{k-1}^{+}, \boldsymbol{\Sigma}_{k-1}^{+})$

➢ 状态方程：$\boldsymbol{X}_k = \boldsymbol{F}\boldsymbol{X}_{k-1} + \boldsymbol{Q}_k$

➢ 观测方程：$\boldsymbol{Y}_k = \boldsymbol{H}\boldsymbol{X}_k + \boldsymbol{R}_k$

卡尔曼滤波

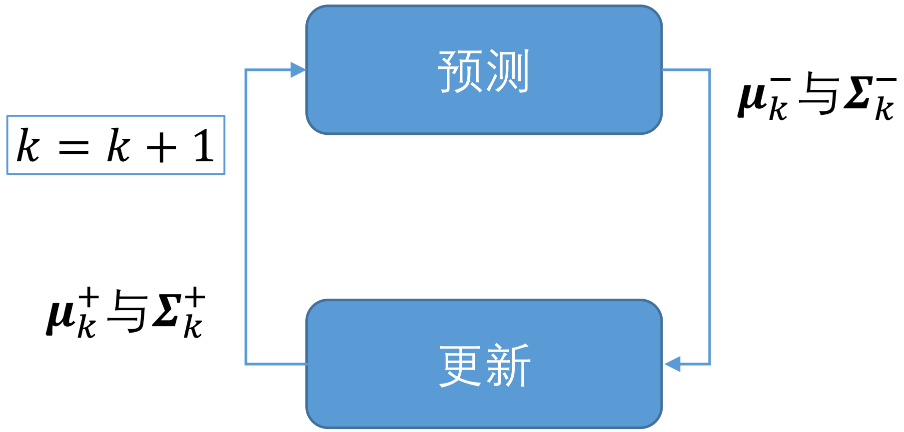
➢ (1) 一步预测状态：$\boldsymbol{\mu}_k^{-} = \boldsymbol{F}\boldsymbol{\mu}_{k-1}^{+}$

➢ (2) 一步预测方差：$\boldsymbol{\Sigma}_k^{-} = \boldsymbol{F}\boldsymbol{\Sigma}_{k-1}^{+}\boldsymbol{F}^{T} + \boldsymbol{\Sigma}_{\boldsymbol{Q}_k}$

➢ (3) 更新卡尔曼增益：$\boldsymbol{K}_k = \boldsymbol{\Sigma}_k^{-}\boldsymbol{H}^{\mathrm{T}}\left(\boldsymbol{H}\boldsymbol{\Sigma}_k^{-}\boldsymbol{H}^{T} + \boldsymbol{\Sigma}_{\boldsymbol{R}_k}\right)^{-1}$

➢ (4) 更新状态：$\boldsymbol{\mu}_k^{+} = \boldsymbol{\mu}_k^{-} + \boldsymbol{K}_k(\boldsymbol{y}_k - \boldsymbol{H}\boldsymbol{\mu}_k^{-})$

➢ (5) 更新方差：$\boldsymbol{\Sigma}_k^{+} = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H})\boldsymbol{\Sigma}_k^{-}$



- **“最优估计”的理解**

□ 什么是估计?

$X(t) \longrightarrow$ **待求系统状态**

$Z(t) = h[X(t)] + V(t) \xrightarrow{\text{解算}} \hat{X}(t)$ **X(t)的估计**

测量得出的数据　**测量噪声**

□ 估计的本质：从含有噪声的量测信息中提取出有用的信息

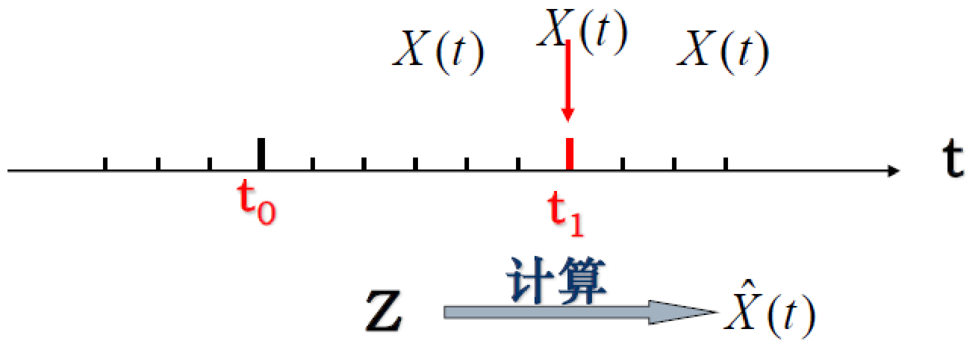
□ 估计的三种表现形式

设在[t_0，t_1]时间段内量测为Z，待求状态为$X(t)$

当$t=t_1$**时,** $\hat{X}(t)$**称为X(t)的估计;**

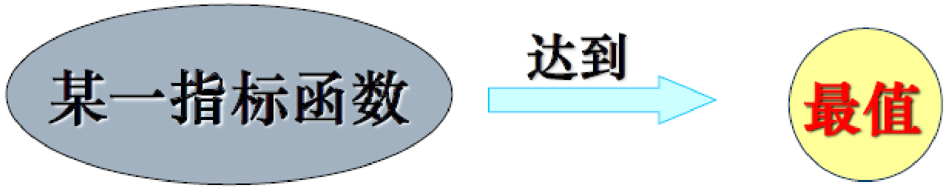
当$t>t_1$**时,** $\hat{X}(t)$**称为X(t)的预测;**

当$t<t_1$**时,** $\hat{X}(t)$**称为X(t)的平滑。**



- **“最优估计”的理解**

□ 什么是最优估计?



则所得估计为最优估计!

□ 估计的本质：从含有噪声的量测信息中提取出有用的信息

□ 估计准则

如上所述,所谓估计问题,就是要构造一个观测数据 Z 的函数 $X(Z)$ 来作为被估计量 $X(t)$ 的一个估计量。我们总希望估计出来的参数或状态变量愈接近实际值愈好。为了衡量估计的好坏,必须要有一个衡量的标准,这个衡量标准就是估计准则。估计常常是以“使估计的性能指标达到极值”作为准则的。估计准则可以是多种多样的。常用的估计准则有:最小方差准则、极大似然准则、极大验后准则、线性最小方差准则、最小二乘准则等。

● “最优估计”的理解

□ 估计准则1：最小二乘估计及其推广

使各量测量 Z_i 与由估计 X 确定的量测估计

$\hat{Z}_i = H_i\hat{X} + V_i$ **之差的平方和最小，即：**

$$\underbrace{J(\hat{X}) = (Z - H\hat{X})^T(Z - H\hat{X}) = \min}$$

$$\Downarrow$$

$$\frac{\partial J}{\partial X}\bigg|_{X=\hat{X}} = -2H^T(Z - H\hat{X}) = 0$$

若H有最大秩n，即H^TH正定，且 $m = \sum_{i=1}^{r} m_i > n$**，则**

$$\hat{X} = (H^TH)^{-1}H^TZ$$

线性估计

特点：

- 最小化量测误差的平方和
- 兼顾所有方程的近似程度，使整体误差最小，有利于抑制噪声
- 无偏估计

假设：量测噪声V是零均值噪声，方差为R的随机向量

（1）最小二乘估计是无偏估计

$E[\hat{X}] = X$或

$E[\tilde{X}] = 0$，式中$\tilde{X} = X - \hat{X}$，为X的估计误差

（2）最小二乘的均方误差阵

$$E[\tilde{X}\tilde{X}^T] = (H^TH)^{-1}H^TRH(H^TH)^{-1}$$

● “最优估计”的理解

□ 估计准则1：最小二乘估计及其推广

□ **加权最小二乘的指标函数**

正定加权矩阵

$$\underbrace{J(\hat{X}) = (Z - H\hat{X})^T W (Z - H\hat{X}) = \min}$$

成立，应满足

$$\frac{\partial J}{\partial X}\bigg|_{X=\hat{X}} = -H^T(W + W^T)(Z - H\hat{X}) = 0$$

$$\hat{X} = [H^T(W + W^T)H)]^{-1} H^T(W + W^T)Z$$

通常W取为对称阵，即$W^T=W$，则加权最小二乘估计为：

$$\hat{X} = [H^T(W)H)]^{-1} H^T(W)Z$$

特点：

➢ 加权最小二乘也是无偏估计

➢ 最优加权最小二乘估计

$$\tilde{X} = X - \hat{X} = -(H^T W H)^{-1} H^T W V$$

量测误差

若V的均值为零，方差阵为R，则：

加权最小二乘估计也是无偏估计！

估计的均方差为：

$$E(\tilde{X}\tilde{X}^T) = (H^T W H)^{-1} H^T W R W H (H^T W H)^{-1}$$

若$W=R^{-1}$，则加权最小二乘估计为：

$$\hat{X} = (H^T R^{-1} H)^{-1} H^T R^{-1} Z$$

马尔可夫估计

马尔可夫估计的均方误差为：

$$E(\tilde{X}\tilde{X}^T) = (H^T R^{-1} H)^{-1}$$

□ **比其他加权最小二乘估计均方差都小**

□ **最优的加权最小二乘估计**

- **“最优估计”的理解**

□ 估计准则1：最小二乘估计及其推广

□ **由前k次量测值确定的加权最小二乘估计为：**

$$\hat{X}_k = \boxed{(\overline{H}_k^T \overline{W}_k \overline{H}_k)^{-1}} \overline{H}_k^T \overline{W}_k \overline{Z}_k$$

令 P_k

□ **由前k+1次量测值确定的加权最小二乘估计为：**

$$\hat{X}_{k+1} = (\overline{H}_{k+1}^T \overline{W}_{k+1} \overline{H}_{k+1})^{-1} \overline{H}_{k+1}^T \overline{W}_{k+1} \overline{Z}_{k+1}$$

推导得：

$$P_{k+1} = P_k - P_k H_{k+1}^T \left(W_{k+1}^{-1} + H_{k+1} P_k H_{k+1}^T\right)^{-1} H_{k+1} P_k$$

$$\hat{X}_{k+1} = \hat{X}_k + P_{k+1} H_{k+1}^T W_{k+1} (Z_{k+1} - H_{k+1} \hat{X}_k)$$

特点：

- 采用批处理实现的最小二乘算法，需存储所有量测值
- 递推最小二乘估计从每次获得的量测值中提取出被估计量的信息，用于修正上一步所得的估计
- 获得量测的次数越多，修正的次数也越多，估计精度也越高

□ **只要给定初始值X_0和P_0，就能获得X在任意时刻的最小二乘估计；**

□ **初始值X_0和P_0可以任意选取，随着量测次数的增加，初值影响逐渐消失，估计值逐渐趋于稳定而逼近被估计量。**

● “最优估计”的理解

□ 估计准则1：最小二乘估计及其推广

类别：

- 最小二乘估计
- 加权最小二乘估计
- 最优加权最小二乘估计
- 递推最小二乘估计

□ 最小二乘估计优、缺点

优点：

算法简单，特别是一般最小二乘估计，不必知道与被估计量及量测量有关的任何统计信息。

缺点：

- □ **只能估计确定性的常值向量，而无法估计随机向量的时间过程；**
- □ **最优指标只保证了量测的估计均方误差之和最小，而并未确保被估计量的估计误差达到最佳，故估计精度不高。**

$$\boxed{J(\hat{X}) = (Z - H\hat{X})^T W (Z - H\hat{X}) = \min}$$

● “最优估计”的理解

□ 估计准则2：最小方差估计及其推广

□ 最小方差估计的估计准则是估计的均方误差最小，即：

Z是m维随机量测向量

$$E\{[X-\hat{X}(Z)]^T[X-\hat{X}(Z)]\} \le E\{[X-\Gamma(Z)]^T[X-\Gamma(Z)]\}$$

估计均方差阵

根据其他方法用Z计算得到的X的估值

□ 最小方差估计的误差小于等于其他估计的均方误差!

特点：

- 无偏估计
- 最小方差即为估计误差的方差
- 通常比较难计算，

□ 最小方差估计具有无偏性质，即它的估计误差（亦可用 $\tilde{X}$ 表示）的均值为零。即：

$$E\{[X-\hat{X}(Z)]\}=E\{\tilde{X}\}=0$$

□ 估计的均方误差就是估计误差的方差，即：

$$E\{\tilde{X}\tilde{X}^T\}=E\{[\tilde{X}-E(\tilde{X})][\tilde{X}-E(\tilde{X})]^T\}$$

□ 因此，最小方差估计不但使估值 $\hat{X}(Z)$ 的均方误差最小，而且这种最小的均方误差就是估计的误差方差

- **“最优估计”的理解**

□ 估计准则2：最小方差估计及其推广

□ **如果将估值 $\hat{X}$ 规定为量测矢量Z的线性函数，即**

$$\hat{X} = AZ + b$$

□ **式中A和b分别是（n×m）阶和n维的矩阵和矢量。这 样的估计方法称为线性最小方差估计。**

□ **可证明，这种估计只需要被估计值X和量测值Z的一、二阶统计特性，所以，它比最小方差估计较为实用。**

特点：

➢ 实用性较高，可以计算

- **“最优估计”的理解**

  - 估计准则3：概率相关估计

    - **贝叶斯估计**

      ——使贝叶斯风险达到最小的估计

    - **极大似然估计**

      ——使关于条件概率密度的似然函数达到极大的估计

    - **极大验后估计**

      ——使验后概率密度函数达到极大的估计

05 卡尔曼滤波(理解与实践)

● “最优估计”的理解

□ 估计准则对比

□ 最小二乘估计

适用于对常值向量或随机向量的估计。

优点：算法简单，对被估计量和量测误差缺乏了解时仍能适用。

缺点：最优指标为量测估计精度达到最佳，估计中不使用被估计量和量测误差的统计信息，估计精度不高。

□ 最小方差估计

优点：所有估计中估计的均方误差为最小的估计，是所有估计中的最佳者。

缺点：只确定估计值是被估计量在量测空间上的条件均值这一抽象关系。而按条件均值的一般求法求取最小方差估计非常困难。

□ 线性最小方差估计

适用于随机过程的估计。

优点：所有线性估计中的最优者，估计过程中只需知道被估计量和量测量的一阶和二阶距。

缺点：平稳过程的一阶和二阶距都为常值，而非平稳过程的一阶和二阶距随时间变化，必须知道每一估计时刻的一阶和二阶距才能求出估计值，这一要求太苛刻。故线性最小方差估计适用于平稳过程而难以适用非平稳过程。

□ 极大验后估计、贝叶斯估计、极大似然估计

与条件概率密度有关，计算困难。

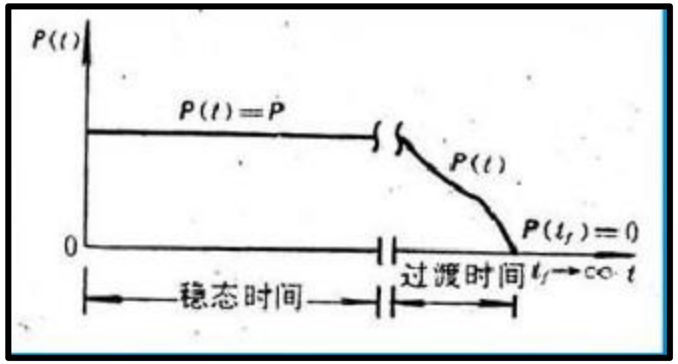
特点:

- 可实现，估计方差小：线性最小方差估计
- 实时性：递推

卡尔曼滤波

- **卡尔曼滤波理解**

□ 连续卡尔曼滤波



特点:

- 微分黎卡提方程的解P(在远离无限时域内稳定，初值扰动的影响较小，很快收敛)，对卡尔曼增益K的响应较小
- Q通过影响P来影响K，R直接影响K
- QR^{-1}正相关K，增大导致卡尔曼滤波倾向于观测值，减小导致卡尔曼滤波倾向于预测值
- R一般由传感器获取，Q是可调参数

□ **但也可以直接从连续系统的滤波方程求出系统状态的估值，设连续系统的状态方程和量测方程分别为：**

$$\dot{X}(t)=F(t)X(t)+G(t)W(t)$$
$$Z(t)=H(t)X(t)+V(t)$$

□ X (t)和Z (t)分别为t时刻的系统状态矢量和量测矢量;

□ F(t)，G(t)和H(t)，分别为系统矩阵，系统噪声矩阵和量测矩阵;

□ W(t)和V(t) 和分别为系统噪声矢量和量测噪声矢量，卡尔曼滤波要求它们都是零均值的白噪声过程。

□ **根据估计均方误差最小的估计准则，按上述系统和量测值，可以推导出连续系统的滤波方程，即**

$$\begin{cases} K(t)=P(t)H^T(t)R^{-1}(t) \\ \dot{\hat{X}}(t)=F(t)\hat{X}(t)+K(t)[Z(t)-H(t)\hat{X}(t)] \\ \dot{P}(t)=P(t)F^T(t)+F(t)P(t)-P(t)H^T(t)R^{-1}(t)H(t)P(t)+G(t)Q(t)G^T(t) \end{cases}$$

□ **R^{-1}(t)为是量测噪声的方差强度阵 R(t)的逆阵，即**

$$E\{V(t)V^T(\tau)\}=R(t)\delta(t-\tau)$$

□ **Q(t)是系统噪声的方差强度阵，即**

$$E\{[G(t)W(t)][G(\tau)W(\tau)]^T\}=G(t)Q(t)G^T(t)\delta(t-\tau)$$

- **卡尔曼滤波理解**

□ 参数稳定性

特点:

➢ 协方差正定性的保证

$$
\begin{aligned}
P_k^+ &= (I - K_k H_k) P_k^- (I - K_k H_k)^{\mathrm{T}} + K_k R_k K_k^{\mathrm{T}} \\
&= [(P_k^-)^{-1} + H_k^{\mathrm{T}} R_k^{-1} H_k]^{-1} \\
&= (I - K_k H_k) P_k^- \qquad (5.19)
\end{aligned}
$$

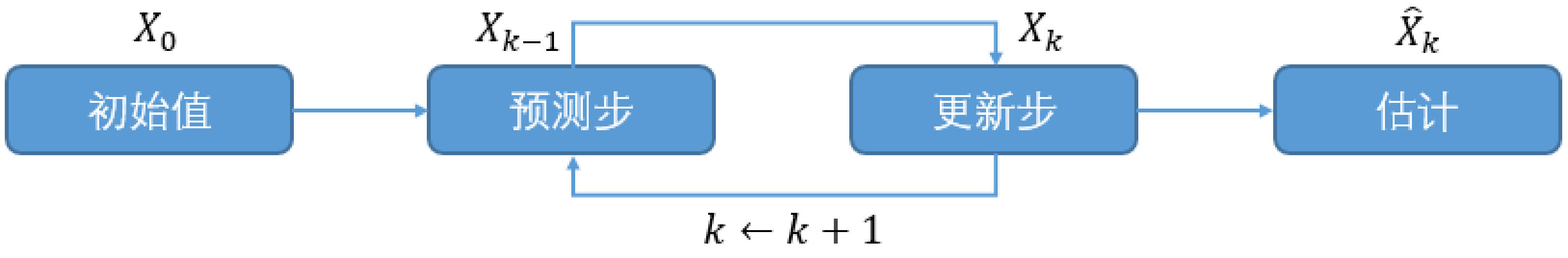
上面 P_k^+ 的第一个表达式叫做协方差量测更新方程的 Joseph 稳定形式。它是 Peter Joseph 在 1946 年创立的,比 P_k^+ 的第三个表达式表现得更稳定、鲁棒性更好[Buc68, Cra04](见习题 5.2)。只要 P_k^- 是对称的正定阵,那么 P_k^+ 的第一个表达式保证了 P_k^+ 将一直是对称的正定阵。P_k^+ 的第三个表达式比第一个表达式计算简单,但是它的形式不能确保 P_k^+ 是对称的或是正定阵。P_k^+ 的第二个表达式很少用到,但是它将用于 6.2 节中信息滤波器的推导。

- **“数据融合”的理解**

□ 预测与量测的数据融合

特点：

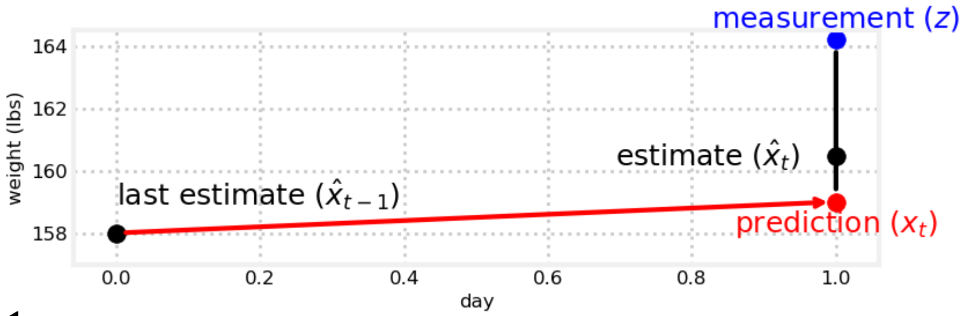
- 多个数据比单个数据要好，不要剔除所谓的低精度数据：将**不准确的测量与不准确的系统模型混合**得到的估计比信息源本身更好
- 预测：基于上一步估计值与系统模型
- 更新：基于预测值与观测值做融合(估计处于两者之间)，融合比例取决于两者的精度(方差)
- 估计值：用期望描述



05 卡尔曼滤波(理解与实践)

● **g-h滤波**

□ Example: 体重

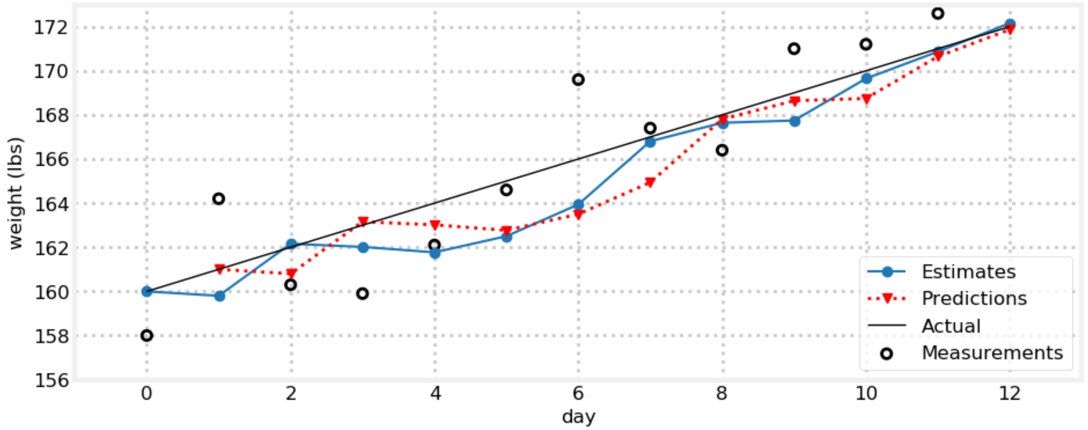


融合增益： $g = 0.4$

$$\text{estimate} = \text{prediction} + \frac{4}{10}(\text{measurement} - \text{prediction})$$

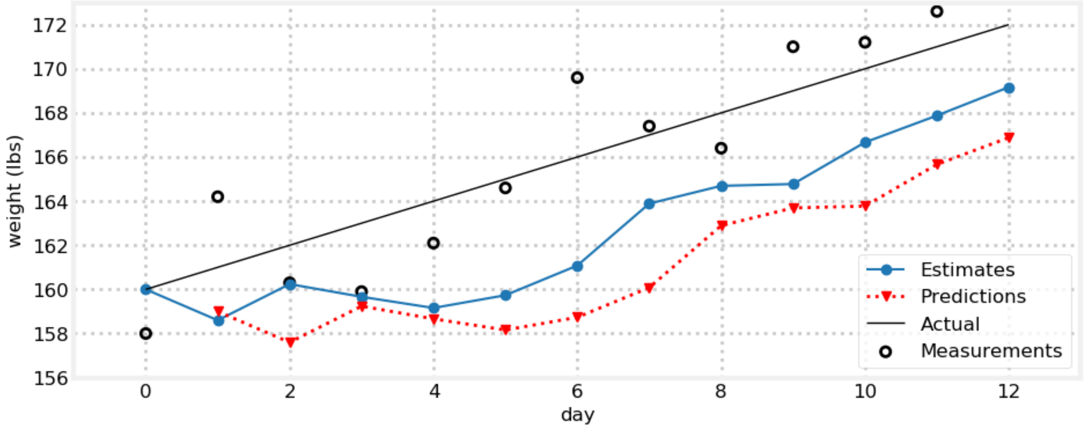
实际体重模型： $dx = 1$

方案1：预测模型： $dx = 1$，每天增加1lb



结论：当预测模型较准时，滤波效果较好，估计值能够很好的跟踪真实值

方案2：预测模型： $dx = -1$，每天减小1lb

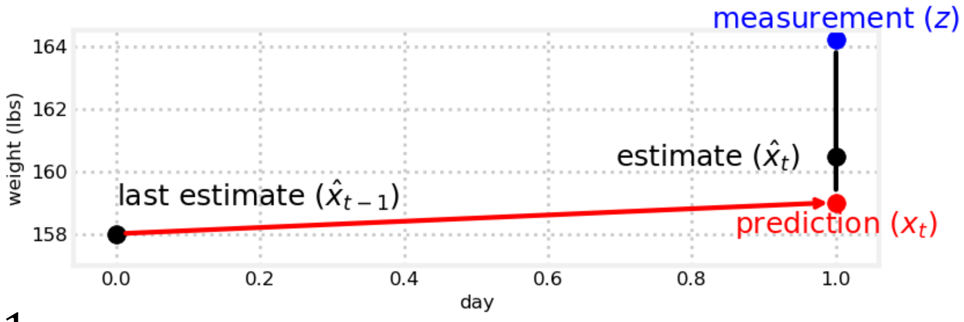


结论：当预测模型很不准时，滤波效果很差，估计值虽然能够有跟踪趋势，但跟踪速度很慢

05 卡尔曼滤波(理解与实践)

- **g-h滤波**

□ Example: 体重



融合增益：$g = 0.4$

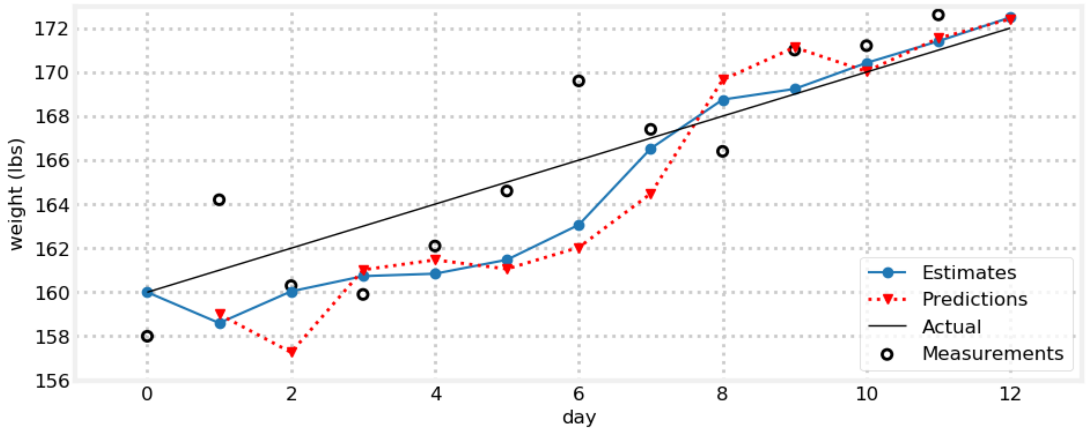
$$\text{estimate} = \text{prediction} + \frac{4}{10}(\text{measurement} - \text{prediction})$$

实际体重模型：$dx = 1$

方案3：利用观测与估计值修正预测模型

$$\text{new gain} = \text{old gain} + \frac{1}{3}\frac{\text{measurement - predicted weight}}{1 \text{ day}}$$

修正增益：$h = 0.4$

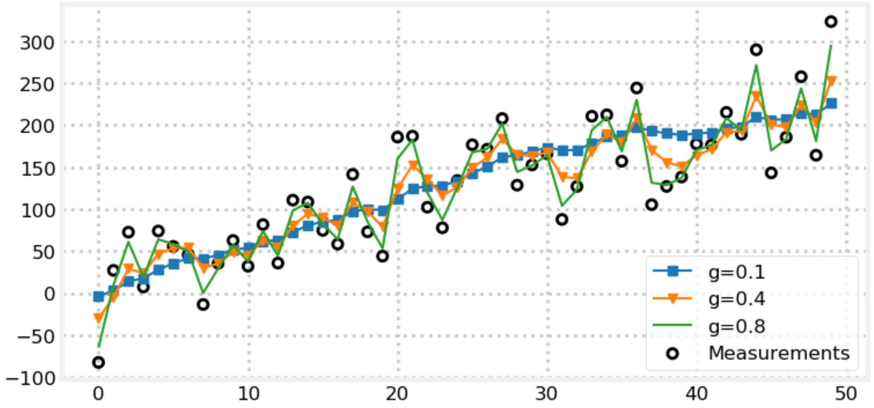


结论：初始模型不正确导致前期误差较大，修正预测模型后后期误差较小

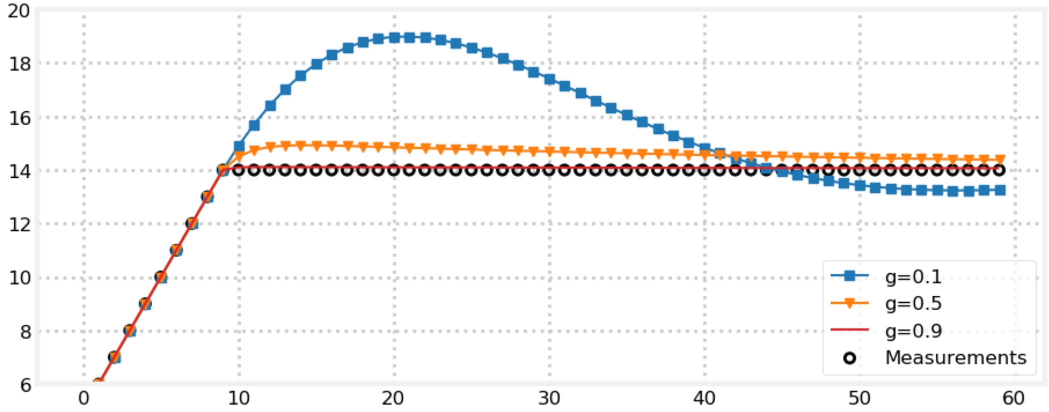
- **g-h滤波**

□ Example: 变增益g

变增益g，增加



变增益g，减小



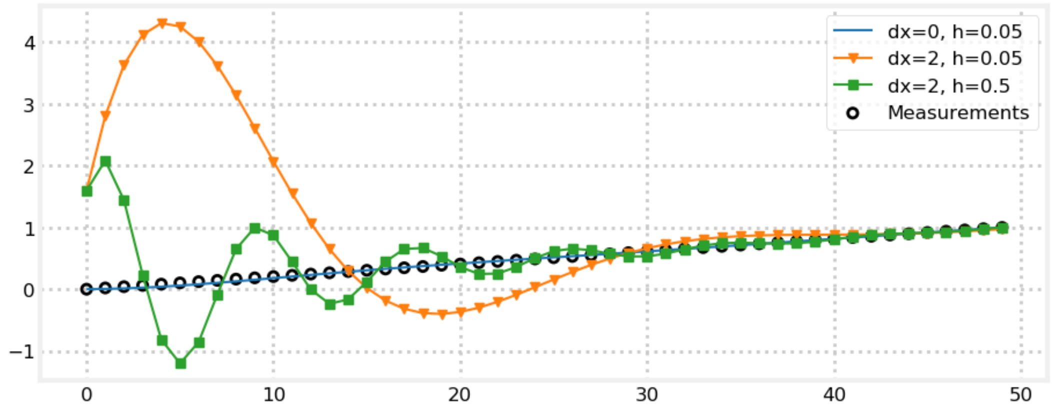
结论：增益g越大，越接近于量测值，滤波结果几乎等同于量测而没有任何噪声的消除作用

结论：增益g越大，越接近于预测值，可以滤掉噪声，但信号变化也无法跟踪上

- **g-h滤波**

  - Example: 变增益h

变增益h，是对状态的预测与量测变化率的权衡，h很大时能够跟踪这种变化，反之不行



结论：增益h越大，跟踪变换能力越强(超调小)，但随着震荡也会变大