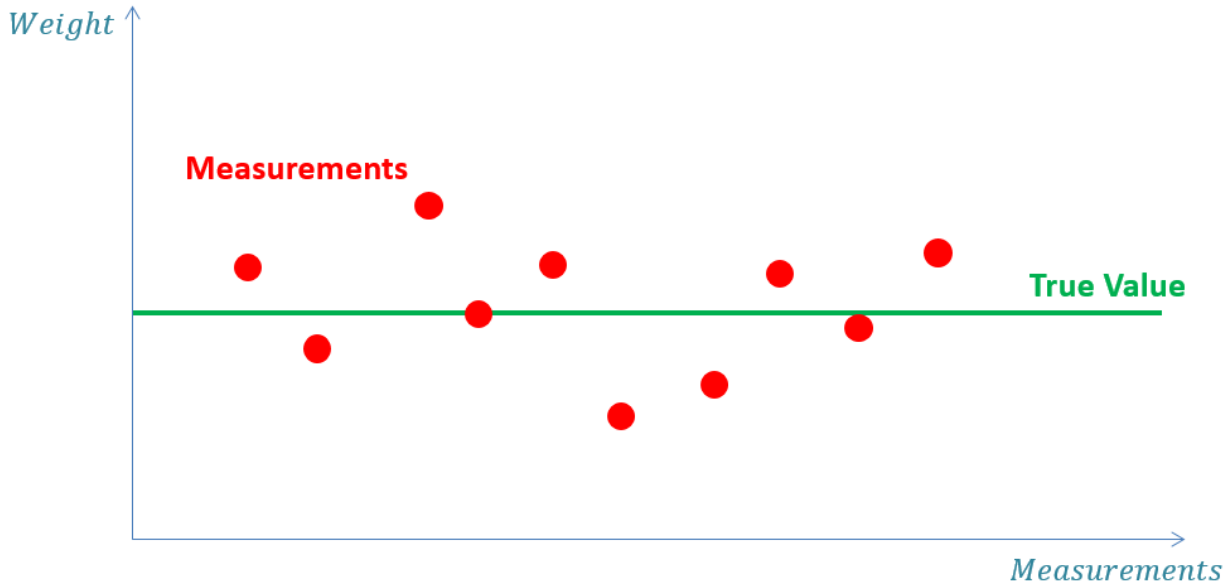
KalmanFilter

以一个例子开头：



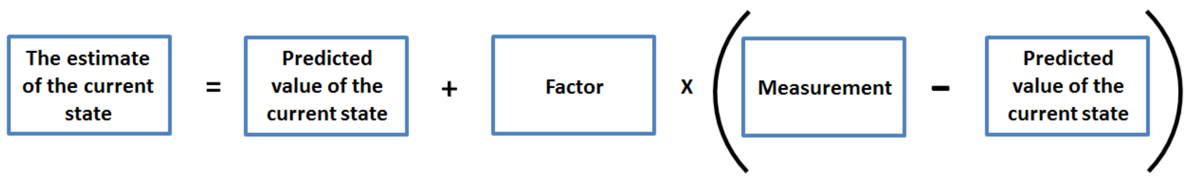
假设我们需要对一个机器人的重量进行估计，我们可以通过机器人的大小，材质，以及实际的测量对一个机器人的重量进行估计，由于大小材质的估计难以准确，以及测量仪器的精度有限，我们无法得知真实的重量，只能得到一个自己的估计值以及测量值。经过多次测量我们得到如上的结果图，面对这种情况，我们为了获得一个更加精确的值，通常可以通过多次测量并求平均值来提高测量的准确性。

$$\hat{X}_n = \frac{1}{n}(Z_1 + Z_2 + \cdots + Z_{n-1} + Z_n) = Average(\sum_{i=1}^{n} Z_i)$$

假如现在又获得了一个测量数据，为了利用上这个数据，我们可以使用如下的公式来更新估计值：

$$\hat{X}_{n+1} = \frac{Z_1 + \cdots + Z_n + Z_{n+1}}{n+1} = \frac{n\hat{X}_n + Z_{n+1}}{n+1} = \hat{X}_n + \frac{1}{n+1}(Z_{n+1} - \hat{X}_n)$$

也就是:



我们可以拓展一下整个过程:
估计:

- 我们先看到这个机器人,我们对重量有一个大致的估计值 $\hat{X}_0$
- 估计我们的估计值的误差为 σ_0

更新:

- 我们可以测量一次这个机器人的重量，得到测量值 Z_1
- 我们可以根据测量值 Z_1 更新我们的估计值 $\hat{X}_1$ (*求平均*)
- 我们可以根据测量值 Z_1 更新我们的估计值的误差 σ_1

不断重复更新

> 上面的过程只包含一次估计,然后就是不断更新,这其实是特例,或者也可以看成是我们在估计过程前后,我们的估计值没有发生变化,也就是估计过程是 $y = x$,估计值没有变化

或者我们可以这样想：
初始化:

- 假设我们先只看到了机器人，先对重量有个大致的估计值 $\hat{X}_0$
- 估计我们的估计值的误差为 σ_0

估计:

- 我们摸了摸机器人，大致估计了部分机器人的材质，对重量有个更加精确的估计值 $\hat{X}_1^-$
- 估计我们的估计值的误差为 σ_1^-

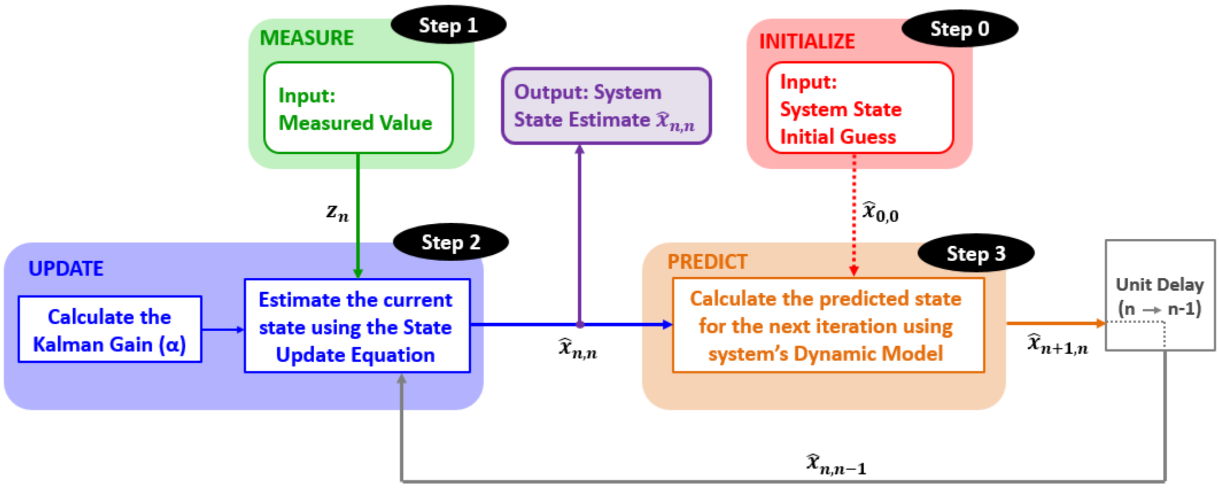
更新:

- 我们又称了称机器人，得到测量值，根据测量值 Z_1 更新估计值 $\hat{X}_1$
- 我们可以根据测量值 Z_1 更新我们的估计值的误差 σ_1

不断重复 估计,更新

> 这样的话，我们的估计过程的估计值就是不断在更新的，也就是估计过程是 $y = f(x)$,估计值在不断变化
> 更新过程是测量值和估计值的比较

用图来全面展示一下:



其实这就是我们的kalman filter的过程，我们先对状态进行估计，然后通过测量值对估计值进行更新，然后不断重复这个过程，这样我们就可以得到一个更加精确的估计值。

kalman filter 黄金五式

预测：

$$\text{先验估计：}\hat{X}_k^- = F\hat{X}_{k-1} + BU_{k-1}$$
$$\text{先验误差协方差：}P_k^- = FP_{k-1}F^T + Q$$

更新：

$$\text{卡尔曼增益：}K_k = \frac{P_k^- H^T}{HP_k^- H^T + R}$$
$$\text{后验估计：}\hat{X}_k = \hat{X}_k^- + K_k(Z_k - H\hat{X}_k^-)$$
$$\text{更新误差协方差：}P_k = (I - K_k H)P_k^-$$

> 初值：$\hat{X}_0$，P_0
> F：状态转移矩阵，B：控制矩阵，U：控制向量，Q：过程噪声协方差矩阵，H：观测矩阵，R：观测噪声协方差矩阵，I：单位矩阵

协方差矩阵： [参考bilibili Dr_can]

> 协方差矩阵是一个用来描述**多个变量**之间关系的矩阵。你可以把它想象成一个表格，每一行和每一列都代表一个变量，每个单元格里的数字就是两个变量之间的协方差。协方差是一个衡量两个变量的相似程度的指标，它的正负和大小都有意义。如果协方差是正的，说明两个变量是正相关的，也就是说，一个变量增大时，另一个变量也会增大；如果协方差是负的，说明两个变量是负相关的，也就是说，一个变量增大时，另一个变量会减小；如果协方差是零，说明两个变量没有线性关系，也就是说，一个变量的变化不会影响另一个变量的变化。协方差的绝对值越大，说明两个变量的相关程度越强，反之则越弱。

举个例子:
假设我们有三个变量，分别是学生的数学成绩 X 、英语成绩 Y 和体育成绩 Z ，我们可以用一个3×3的协方差矩阵来表示它们之间的关系，如下：

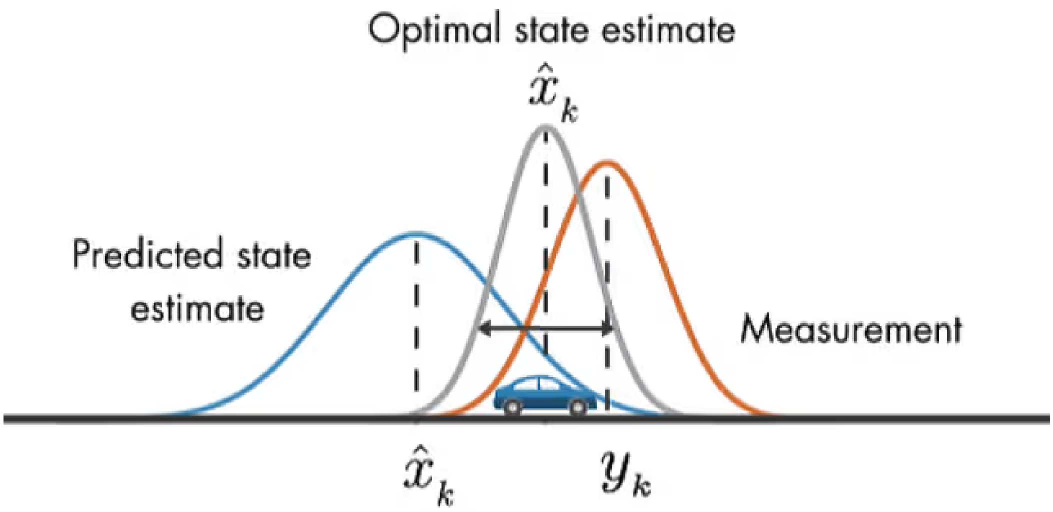
$$\begin{bmatrix} cov(X,X) & cov(X,Y) & cov(X,Z) \\ cov(Y,X) & cov(Y,Y) & cov(Y,Z) \\ cov(Z,X) & cov(Z,Y) & cov(Z,Z) \end{bmatrix}$$

- $cov(X, X)$ 就是数学成绩和数学成绩之间的协方差，也就是数学成绩的方差，它的值越大，说明学生的数学成绩越分散，也就是说，学生的数学成绩差异越大；
- $cov(X, Y)$ 就是数学成绩和英语成绩之间的协方差，它的值越大，说明学生的数学成绩和英语成绩之间的相关性越强，也就是说，学生的数学成绩和英语成绩越相似；
- $cov(X, Z)$ 就是数学成绩和体育成绩之间的协方差，它的值越大，说明学生的数学成绩和体育成绩之间的相关性越强，也就是说，学生的数学成绩和体育成绩越相似；以此类推。

其实kalman filter的过程就是一个数据融合的过程,我们把我们对这个过程的先验知识和测量值进行融合，得到一个更加精确的估计值。
因为物理真值不可能获得，所以我们只能通过测量值来估计物理真值，但是测量值又不是完全可靠的，所以我们需要对测量值进行修正，这就是kalman filter的过程。
Data fusion

Z_1服从a, σ_1的正太分布, Z_2服从b, σ_2的正太分布

估计真实取值：$\hat{Z} = Z_1 + K(Z_2 - Z_1)$，求得合适的K使得$\hat{Z}$的波动最小，即方差最小。



详情参考 Dr_can https://www.bilibili.com/video/BV12D4y1S7fU/?spm_id_from=333.788&vd_source=189e765254c924184e8a13b2384d02d2

以一个例子继续：
假设我们在估计一辆在平地上运动的车辆的状态，状态为$x = [s, v]^T$, s为车辆的位置，v为车辆的速度，假设我们固定dt测量一次车辆的位置，我们可以得到如下的状态转移方程：

$$x_{k+1} = \begin{bmatrix} 1 & dt \\ 0 & 1 \end{bmatrix} x_k + w_k = Fx_k + Q$$

> 假设我们对汽车运动状态的估计是**匀速直线运动**，那么我们可以得到如下的状态转移方程， w_k代表在第k个dt我们估计的运动状态和实际的误差，比如实际上汽车有加速度

我们用一个卫星测量这个汽车的位置，假设测量误差为v_k，那么我们可以得到如下的测量方程：

$$z_k = [1 \quad 0] x_k + v_k = Hx_k + R$$

我们可以得到如下的**协方差矩阵**：

$$Q = \begin{bmatrix} E(ss^T) & E(sv^T) \\ E(s^T v) & E(vv^T) \end{bmatrix}, R = var(v)$$

这些代表了我们对各个状态之间的相关性的估计，比如我们认为位置和速度之间的相关性为0，那么我们就把他们的协方差值设置为零，通常情况下我们都认为各个状态变量之间相互不相关

$$Q = \begin{bmatrix} {\sigma_s}^2 & 0 \\ 0 & {\sigma_v}^2 \end{bmatrix}, R = var(v)$$

> 也就是各个变量自己的方差

开始计算 具体过程查看excel, 密码是下面这个问题的答案:
假设 $Z_1 = 6.5mm, Z_2 = 7.3mm$, 我们假设已知 Z_1 的测量波动标准差为 $\sigma_1 = 0.2mm$, Z_2 的测量波动标准差为 $\sigma_2 = 0.4mm$，求我们估计得到的最优值应该是多少?

简略推导(啥也没讲)

实际情况state space 方程：

$$X_k = AX_{k-1} + BU_{k-1} + W_{k-1}(\text{过程噪声})$$
$$Z_k = HX_k + V_k(\text{测量噪声})$$

X_k：第k次状态值，U_{k-1}：第$k-1$次的输入，W_{k-1}：第$k-1$次的过程噪声, Z_k：第k次的观测取值，V_k：第k次的测量噪声

假设 **W**，**V** 服从期望为 0 ，协方差矩阵为 **Q** 和 **R** 的正态分布， $cov(A) = E[AA^T]$。

$$\hat{X_k^-} = AX_{k-1} + BU_{k-1}：\text{先验估计取值（建模计算）}$$
$$\hat{X_{kmea}} = H^-Z_k：\text{观测估计取值（测量）}$$

可以通过以下计算得到后验取值

$$\hat{X}_k = \hat{X}_{k-1}^- + G(H^-Z_k - \hat{X}_{k-1}^-)$$
$$G = K_kH$$

目标：取 K_k 使得 $\hat{X}_k$ 的误差最小，即 $e_k = X_k - \hat{X}_k$ 的误差最小 ，e_k 是服从期望为 0 ，协方差矩阵为 **P** 的正太分布，所以就是 $P = E[ee^T]$（协方差的定义式）最小，而根据矩阵论的知识，可以知道$tr(P)$最小就是P最小。

> 根据 P 的定义式推导，带入已经知道的 e ，得到 tr(p) 的表达式，再根据矩阵求导，对 Kk 求导，并令导数为零可以求得最后的 Kk 取值。

经过一系列推导可以求得当 K_k取值如下时，可以获得最小的后验估计误差：

$$K_k = \frac{P_k^-H^T}{HP_k^-H^T + R}$$

其中，R：观测误差的噪声，H：观测矩阵，P_k^-:k时刻的先验误差的协方差矩阵，$e_k^- = X_k - \hat{X}_k^-$:先验误差（先验估计值和实际取值之间的误差）。

$P_k^- = E[e_k^-e_k^{-T}]$，带入e_k^-通过推导，可以得到

$$P_k^- = AE[e_{k-1}e_{k-1}{}^T]A^T + E[W_{k-1}W_{k-1}{}^T]$$
$$\text{即}P_k^- = AP_{k-1}A^T + Q$$

Extended Kalman Filter

Nonlinear Equation：

$$X_k = f(X_{k-1}, U_{k-1}, W_{k-1})$$
$$Z_k = h(X_k, V_k)$$

通过泰勒展开线性化，运用Kalman filter进行状态估计。由于系统存在各种误差，无法得知真实取值点，所以在$\hat{X}_{k-1}$处展开，线性化，即：

$$X_k = f(\hat{X}_{k-1}, U_{k-1}, W_{k-1}) + A(X_k - \hat{X}_{k-1}) + WW_{k-1}$$
$$A = \frac{\partial f}{\partial X}|X = \hat{X}_{k-1}, U = U_{k-1}$$
$$W = \frac{\partial f}{\partial W}|X = \hat{X}_{k-1}, U = U_{k-1}$$
$$Z_k = h(X_k, V_k) + H(X_k)$$

预测：

$$\text{先验计算}：\hat{X_k^-} = f(\hat{X}_{k-1}, U_{k-1}, W_{k-1})$$
$$\text{先验误差协方差}：P_k^- = AP_{k-1}A^T + WQW^T$$

校正：

$$卡尔曼增益：K_k = \frac{P_k^- H^T}{HP_k^- H^T + VRV^T}$$

$$后验估计：\hat{X}_k = \hat{X}_k^- + K_k(Z_k - h(\hat{X}_k^-, V_k))$$

$$更新误差协方差：P_k = (I - K_k H)P_k^-$$