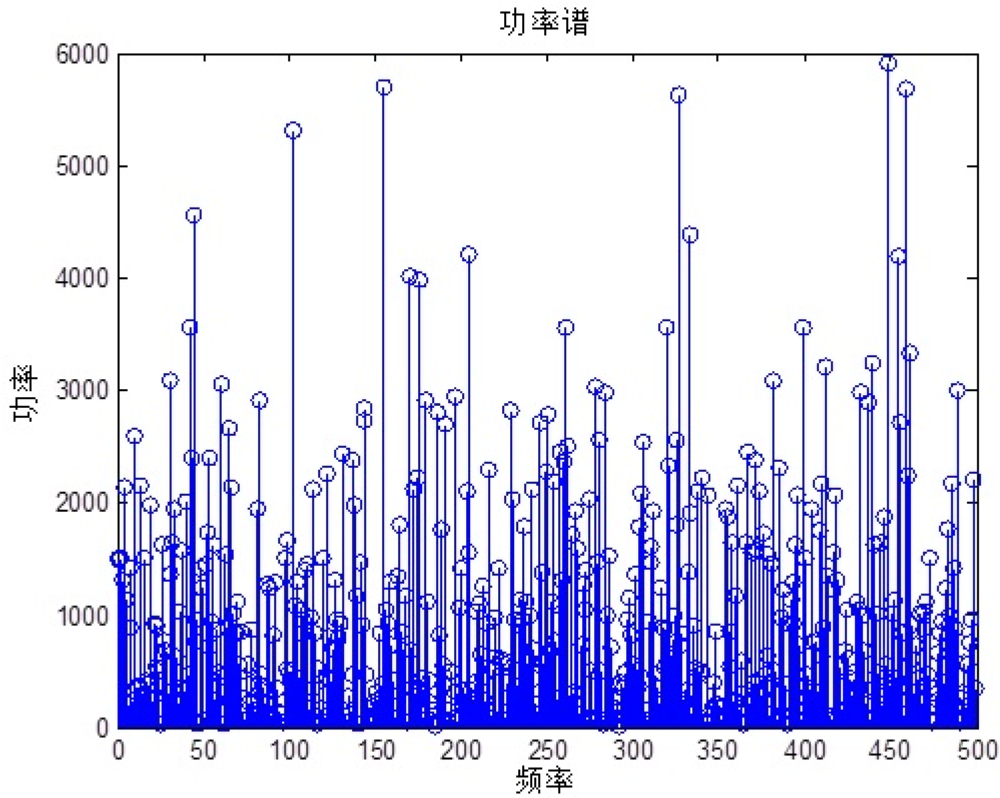
卡尔曼滤波

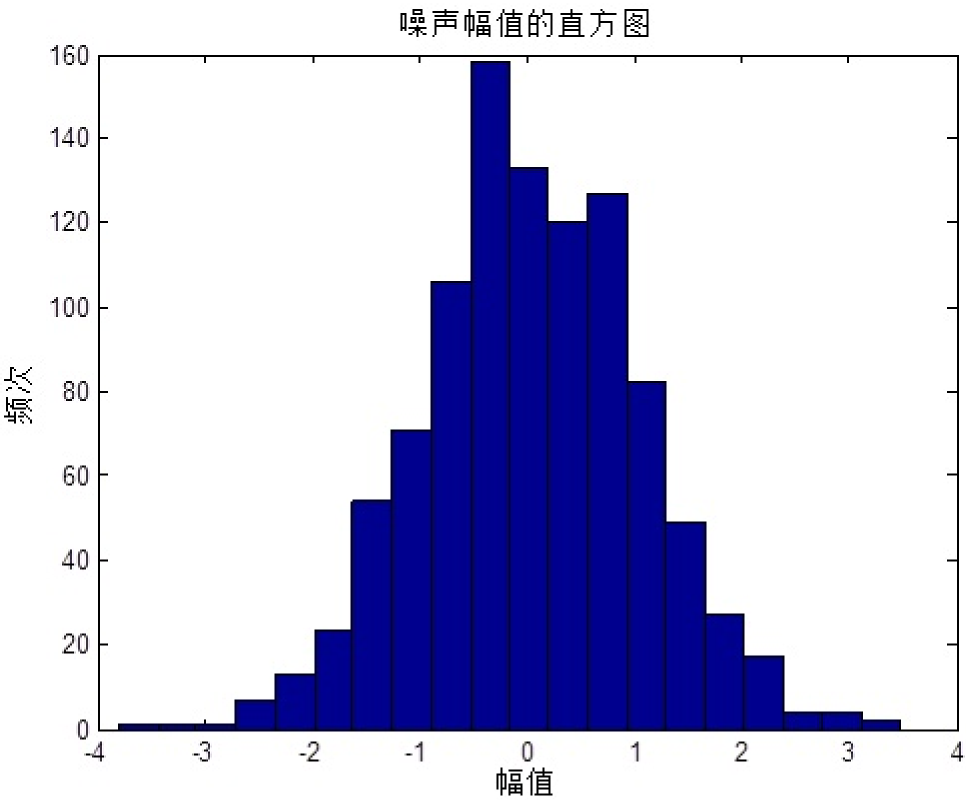
高斯白噪声

> 高斯白噪声，幅度分布服从高斯分布，功率谱密度服从均匀分布

- 白噪声在功率谱（以频率为横轴，信号幅度的平方为功率，功率为纵轴）上趋近于常值，即噪声频率丰富，在整个频谱上都有成分（低频指的是信号不变或缓慢变化，高频是指信号突变）
- 功率谱密度服从均匀分布，就是**白噪声**
  - 功率谱就是频率域的图
  - 由傅里叶变换性质可知，**时域有限，频域无限；频域有限，时域无限**。于是频域无限的信号变换到时域上，对应于冲激函数的整数倍（应该包含0倍）。于是在时间轴的某点上，噪声孤立，与其他点的噪声无关，时间轴上的每个点的幅值都跟其他点无关。简而言之，**任意时刻出现的噪声幅值都是随机的**。
  - 这里所谓的均匀分布应该是大致的均匀分布，每个频率都有值



- 幅度分布服从高斯分布，就是**高斯噪声**
  - 幅度分布指的是时域的图，其中**信号幅值为横轴，出现的频率为纵轴**
  - 这也就是说在一个时间段内，信号出现是在一定范围内的正态分布



- 补充：***相关***是什么意思，作为白噪声信号随机的对立面
    - 信号幅值与前后的信号幅值有关系

    > 例子：此时刻的幅值比上一时刻的大，而下一时刻的噪声幅值比此时刻的还大

协方差矩阵

- 对于共d个随机变量，协方差矩阵为d*d的矩阵，对角线为d个随机变量的方差，其他地方是两两随机变量的协方差
- 协方差矩阵为**对称矩阵**

$$\Sigma=\begin{bmatrix}\sigma(x_1,x_1) & \cdots & \sigma(x_1,x_d)\\ \vdots & \ddots & \vdots\\ \sigma(x_d,x_1) & \cdots & \sigma(x_d,x_d)\end{bmatrix}\in\mathbb{R}^{d\times d}$$

多元正态分布

均值向量为 $\mu=\mu_{n\times 1}=(\mu_1,\mu_2,\ldots,\mu_n)$

协方差矩阵 $\Sigma=\Sigma_{n\times n}$

分布函数为 $p(x;\mu,\Sigma)=\frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$

其中 $(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}=\int_{\mathbb{R}^n}e^{-(\boldsymbol{x}-\boldsymbol{\mu})'(\boldsymbol{\Sigma})^{-1}(\boldsymbol{x}-\boldsymbol{\mu})/2}d\boldsymbol{x}$

以二元正态分布为例：

$\mu=\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}$，$\Sigma=\begin{bmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{bmatrix}$ 显然 $\sigma_{12}=\sigma_{21}$ (同一对随机变量的协方差相等)

$$\Sigma^{-1}=\frac{1}{\sigma_{11}\sigma_{22}-\sigma_{12}^2}\begin{bmatrix}\sigma_{22} & -\sigma_{12}\\ -\sigma_{21} & \sigma_{11}\end{bmatrix}$$

这里我们设相关系数 $\rho_{12}:=\frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}}$

二元正态分布的概率密度函数（pdf）为

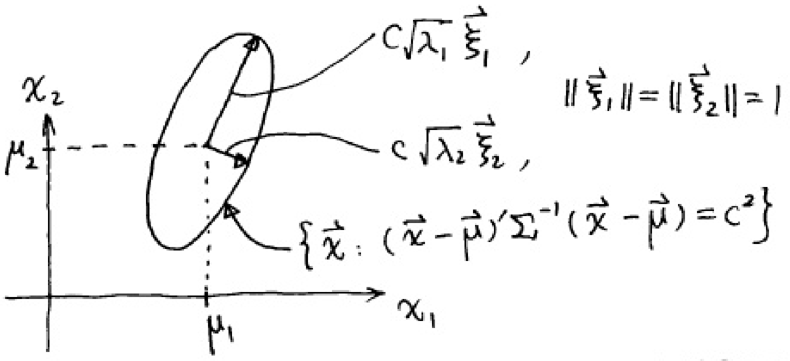
$$f(x_1,x_2)=\frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}\left(1-\rho_{12}^2\right)}}e^{-(\boldsymbol{x}-\boldsymbol{\mu})'(\boldsymbol{\Sigma})^{-1}(\boldsymbol{x}-\boldsymbol{\mu})/2}$$

$$\begin{aligned}
&(x-\mu)'(\Sigma)^{-1}(x-\mu)\\
&=\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\end{bmatrix}'\frac{1}{\sigma_{11}\sigma_{22}-\sigma_{12}^2}\begin{bmatrix}\sigma_{22} & -\rho_{12}\sqrt{\sigma_{11}\sigma_{22}}\\ -\rho_{12}\sqrt{\sigma_{11}\sigma_{22}} & \sigma_{11}\end{bmatrix}\begin{bmatrix}x_1-\mu_1\\ x_2-\mu_2\end{bmatrix}\\
&=\frac{\sigma_{22}(x_1-\mu_1)^2+\sigma_{11}(x_2-\mu_2)^2-2\rho_{12}\sqrt{\sigma_{11}\sigma_{22}}(x_1-\mu_1)(x_2-\mu_2)}{\sigma_{11}\sigma_{22}\left(1-\rho_{12}^2\right)}\\
&=\frac{1}{1-\rho_{12}^2}\left[\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)^2+\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)^2-2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]
\end{aligned}$$

- 对于以二元正态分布为例子，它的投影如下所示：

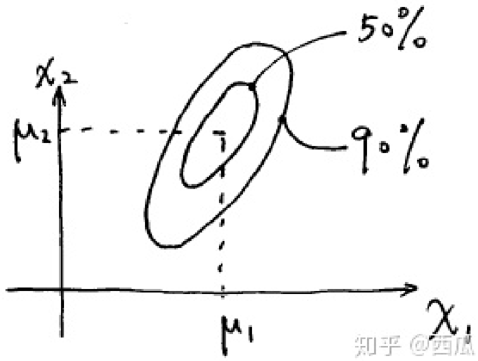
假如把它投影到 z=c 平面上（数学语言为 $\left\{x:(x-\mu)'(\Sigma)^{-1}(x-\mu)=c^2\right\}$）

则投影图像为椭圆，椭圆中心在点 (μ_1,μ_2) ,长短轴为 $c\sqrt{\lambda_i}\boldsymbol{\xi}_i$ 。这里的 ξ_i 为 Σ 矩阵的特征向量，即满足等式 $\Sigma\xi_i=\lambda_i\xi_i$ 的向量解。如下图所示：



我们用 $\chi^2_n(\alpha)$ 表示自由度为n-1 的卡方分布（chi-square distribution）置信度为 α α

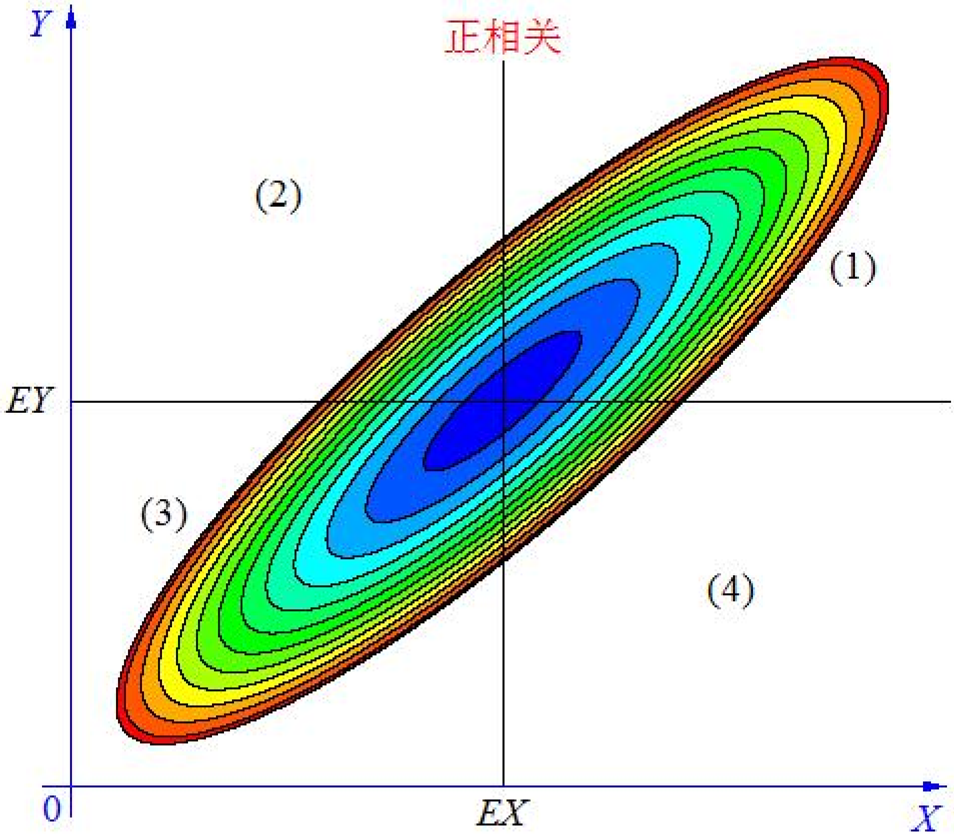
则有 $(\boldsymbol{x}-\boldsymbol{\mu})'(\boldsymbol{\Sigma})^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\leq\chi^2_n(\alpha)$ 的概率为（1- α ）。

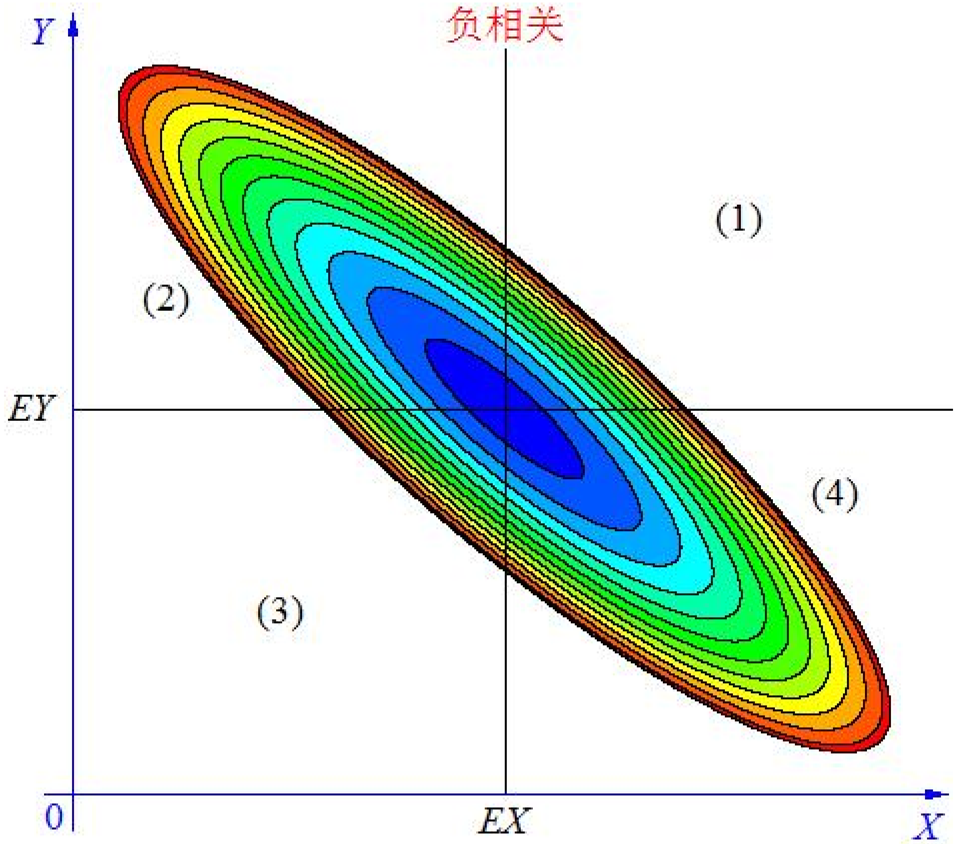


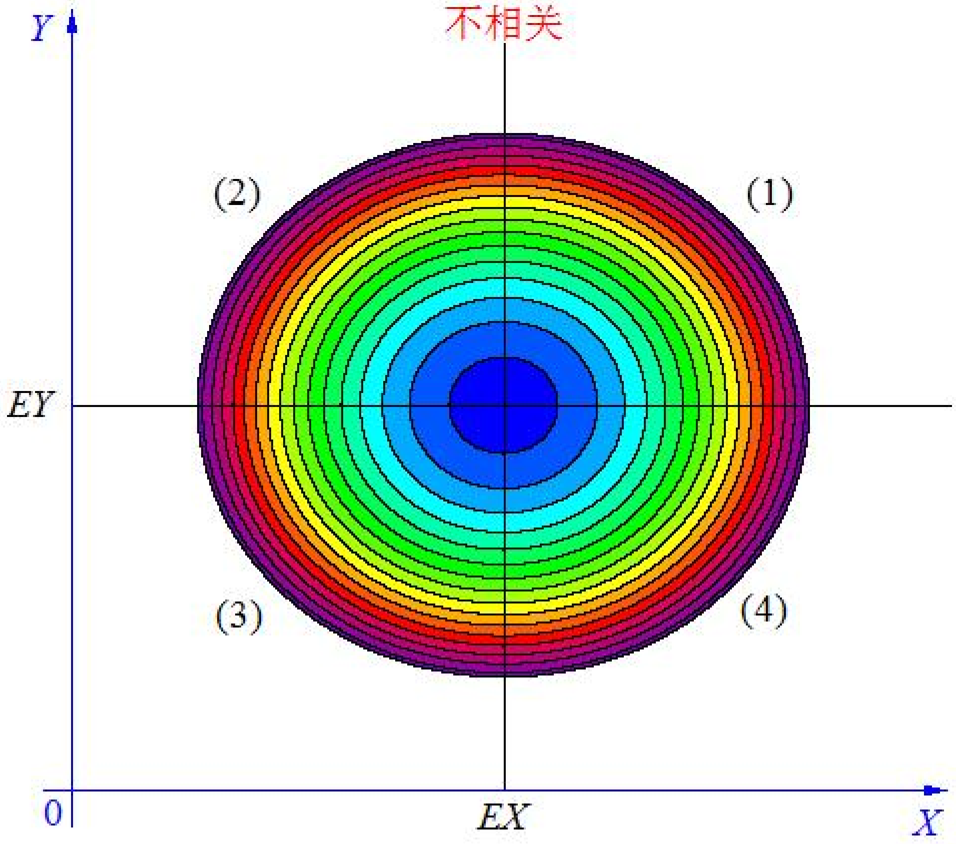
二维正态分布在50%和90%置信度下的等高线轮廓（contours）

协方差矩阵与多元正态函数

- 如上方的图片所示，二元正态函数的投影是一个椭圆，从中可以体现出协方差的意义


正相关
Y
(2)
(1)
EY
(3)
(4)
0
EX
X


负相关
Y
(1)
(2)
EY
(4)
(3)
0
EX
X



- 以此类推其他多元正态函数以及其协方差矩阵，可以感受他们的意义

卡尔曼滤波

> https://www.bzarg.com/p/how-a-kalman-filter-works-in-pictures/

某一时刻

- 对于任意一个系统的任意一个时刻，我们可以用几个量来进行描述。将我们需要的状态参量表达出来，并描述系统的变化情况。这里就有我们的向量

  $$\boldsymbol{x_k}$$

  > 此处类似于物理中的相空间的概念

  一般地，我们有两种方法得到此时刻的下一时刻该系统的状态：

  1. 由此时刻的系统状态对下一时刻系统状态的预测结果
  2. 由传感器等方法在下一时刻得到的结果

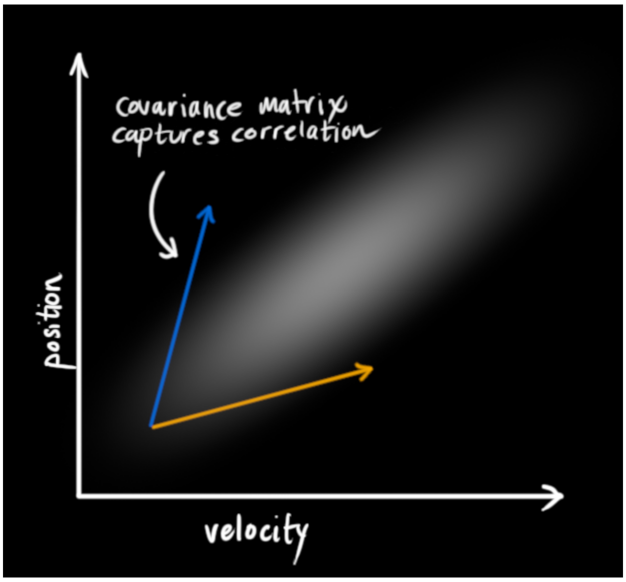
  可是无论通过什么方法得到的系统状态，都不是绝对准确的，其中都包含着**不确定性**。

- 这种不确定性就是**噪声**，我们认为所有这些噪声都是上文所说的**高斯白噪声**，在时域中服从的就是上文所说的多元正态分布。多元正态分布中的

  $$m_i \qquad i \in \{1, 2 \cdots k\}$$

  间都有关系，描述这种关系的就是**协方差矩阵**（在二元正态分布中就是相关系数）

  $$\boldsymbol{P_k}$$



图中所示系统状态量为位置和速度两个量

> 我们为什么需要协方差矩阵呢?
>
> 协方差矩阵将各个描述系统状态的量都联系在一起，相当于增加了信息，让我们可以预测下一时刻的状态

某一时刻到下一时刻

- 对于状态的更新：从现在这一时刻的状态预测下一个时刻的状态我们可以通过一个矩阵来描述

$$\boldsymbol{F}_k$$

- 对于协方差矩阵的更新:

  > 首先介绍一个数学公式
  >
  > $$Cov(x) = \Sigma$$
  > $$Cov(\boldsymbol{A}x) = \boldsymbol{A}\Sigma\boldsymbol{A}^T$$

  于是我们可以更新协方差矩阵了

$$\boldsymbol{P}_k = \boldsymbol{F}_k\boldsymbol{P}_{k-1}\boldsymbol{F}_k^T$$

额外的影响

- 对于实际应用中，上面的更新并不能完全适用。具体还有以下两点:

1. **我们施加的影响**
   - 我们施加的影响的形式是我们已知的，可以用数学表达的，因此较为容易添加
   - 我们施加的影响可以用一个向量表示

$$\boldsymbol{u}_k$$

   - 这个我们施加的影响是通过一个称为**控制矩阵**的东西进行的

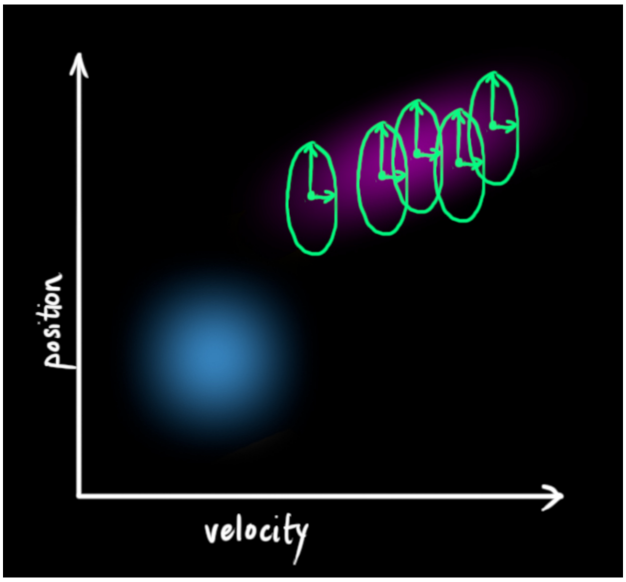
$$\boldsymbol{B}_k$$

   综合这个施加的影响，我们可以更新对下一时刻状态变量的表达式为

$$\boldsymbol{x}_k = \boldsymbol{F}_k\boldsymbol{x}_{k-1} + \boldsymbol{B}_k\boldsymbol{u}_k$$

2. **外界的不确定性**

    - 外界的不确定性是我们未知的，无法控制的，具体而言就是不知道具体的表达式。为了将这种完全的不确定性添加到我们的考虑范围之内，我们通过在每一次预测时添加一些不确定量来实现
    - 我们假定在当前时刻的状态中的每一个状态量会在下一个时刻变到某个范围之内（这个是符合正态分布的，其中自然也包含着协方差矩阵，均值相同）（图中的绿色的椭圆就代表上面说的*某个范围*）



    - 上述协方差矩阵用

$$\boldsymbol{Q}_k$$

综合上面所有的，我们有

$$\boldsymbol{x}_k = \boldsymbol{F}_k \boldsymbol{x}_{k-1} + \boldsymbol{B}_k \boldsymbol{u}_k$$
$$\boldsymbol{P}_k = \boldsymbol{F}_k \boldsymbol{P}_{k-1} \boldsymbol{F}_k^T + \boldsymbol{Q}_k$$

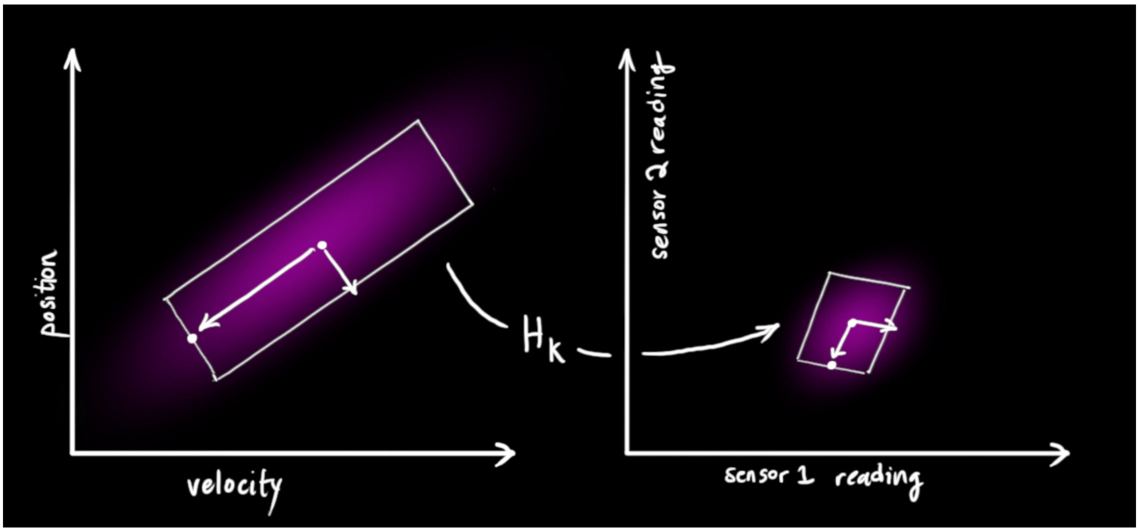
> 用文字表达可以为:
>
> 对下一时刻最好的预测是用这一时刻最好的预测加上一个用已知的额外影响的进行的修正。
>
> 对下一时刻的不确定性的预测是用这一时刻的不确定性加上环境给我们的未知的额外影响

某一时刻自己的更新

- **系统测量矩阵**

  我们通过传感器得到系统现在的状态量和我们一直在跟踪的值(指的是上面的状态量的向量和协方差矩阵)不一定是一模一样的，在此时我们就会需要一个名为系统测量矩阵的东西

$$\boldsymbol{H}_k$$



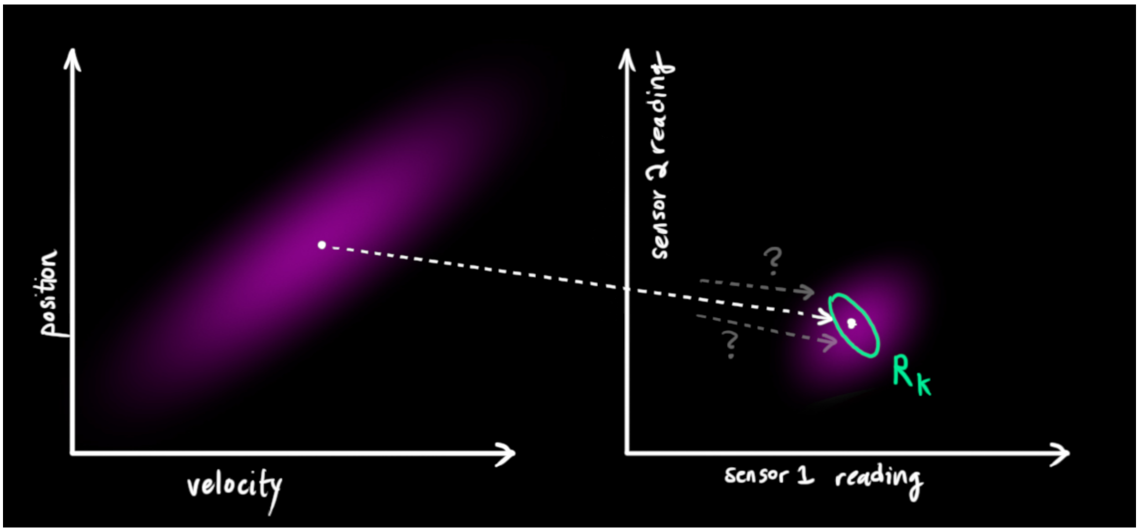
- 在引入了系统测量矩阵之后，我们就有一个更加准确的结果了

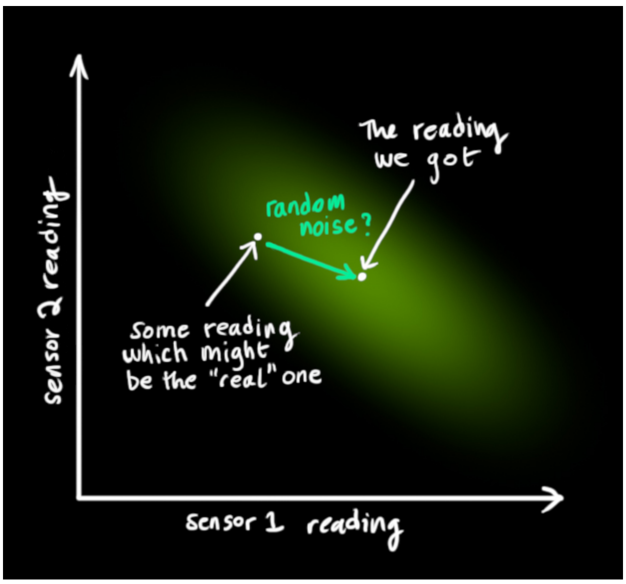
$$\boldsymbol{\mu}_{expected} = \boldsymbol{H}_k \boldsymbol{x}_k$$
$$\Sigma_{expected} = \boldsymbol{H}_k \boldsymbol{P}_k \boldsymbol{H}_k^T$$

- **测量噪声**

众所周知，传感器测量得到的状态量不可能是完全准确的，肯定会有噪声的产生，这个噪声是在系统测量中产生的。同样地，我们认为这个噪声是高斯白噪声，因此有一个协方差矩阵和均值

$$\boldsymbol{R}_k \qquad\qquad \boldsymbol{\mu} = \boldsymbol{z}_k$$





> 我们认为所有都是多元正态分布，描述那个噪声范围的就是协方差矩阵（上文有提及）（图中绿色的椭圆）

- 我们有了系统测量矩阵和测量噪声协方差矩阵。这两个矩阵的分布是不完全相同的，如上面两个图片一样。为了能够将两个结果结合在一起，简单的方法就是将两个正态分布**乘**在一起，只留下重叠的部分。这就是我们能够得到的最好的预估结果。
  - 由数学结论可以得到，两个正态分布的乘积仍然是正态分布

$$\begin{gathered}N(x,\mu_0,\sigma_0)N(x,\mu_1,\sigma_1)=N(x,\mu',\sigma')\\ \mu'=\mu_0+k(\mu_1-\mu_0)\\ \sigma'^2=\sigma_0^2-k\sigma_0^2\end{gathered}$$

    其中

$$k=\frac{\sigma_0^2}{\sigma_0^2+\sigma_1^2}$$

  - 对于矩阵形式，我们可以改写上面的等式为

$$\begin{gathered}\boldsymbol{K}=\Sigma_0(\Sigma_0+\Sigma_1)^{-1}\\ \boldsymbol{\mu}'=\boldsymbol{\mu}_0+\boldsymbol{K}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0)\\ \boldsymbol{\Sigma}'=\boldsymbol{\Sigma}_0-\boldsymbol{K}\boldsymbol{\Sigma}_0\end{gathered}$$

  - 将这个结论具体应用到这个应用场景中，我们可以得到

$$\begin{gathered}(\boldsymbol{\mu}_0,\boldsymbol{\Sigma}_0)=(\boldsymbol{H}_k\boldsymbol{x}_k,\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{H}_k^T)\\ (\boldsymbol{\mu}_1,\boldsymbol{\Sigma}_1)=(\boldsymbol{z}_k,\boldsymbol{R}_k)\end{gathered}$$

  - 在我们得到新的正态分布的时候，我们可以设想到，新的均值和协方差矩阵仍有原来的系统测量时的形式

$$\begin{gathered}\boldsymbol{\mu}'_{expected}=\boldsymbol{H}_k\boldsymbol{x}'_k\\ \Sigma'_{expected}=\boldsymbol{H}_k\boldsymbol{P}'_k\boldsymbol{H}_k^T\end{gathered}$$

- 于是，综上所述，我们有更加准确的结果（下面的结果就是将上面的三段数学式子结合在一起然后消掉了一个H）

$$\begin{gathered}\boldsymbol{x}'_k=\boldsymbol{x}_k+\boldsymbol{K}'(\boldsymbol{z}_k-\boldsymbol{H}_k\boldsymbol{x}_k)\\ \boldsymbol{P}'_k=\boldsymbol{P}_k-\boldsymbol{K}'\boldsymbol{H}_k\boldsymbol{P}_k\\ \boldsymbol{K}'=\boldsymbol{P}_k\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{H}_k^T+\boldsymbol{R}_k)^{-1}\end{gathered}$$

最后的结果

我们有了五个式子组成了卡尔曼滤波的全部公式

$$\begin{gathered}\boldsymbol{x}_k=\boldsymbol{F}_k\boldsymbol{x}_{k-1}+\boldsymbol{B}_k\boldsymbol{u}_k\\ \boldsymbol{P}_k=\boldsymbol{F}_k\boldsymbol{P}_{k-1}\boldsymbol{F}_k^T+\boldsymbol{Q}_k\end{gathered}$$

$$\begin{gathered}\boldsymbol{x}'_k=\boldsymbol{x}_k+\boldsymbol{K}'(\boldsymbol{z}_k-\boldsymbol{H}_k\boldsymbol{x}_k)\\ \boldsymbol{P}'_k=\boldsymbol{P}_k-\boldsymbol{K}'\boldsymbol{H}_k\boldsymbol{P}_k\\ \boldsymbol{K}'=\boldsymbol{P}_k\boldsymbol{H}_k^T(\boldsymbol{H}_k\boldsymbol{P}_k\boldsymbol{H}_k^T+\boldsymbol{R}_k)^{-1}\end{gathered}$$