核函数

- 出发点

  - 如果我的数据有足够多的可利用的信息，那么我可以直接做我喜欢的事了，但是现在如果没有那么多的信息，我可不可以在数学上进行一些投机呢?

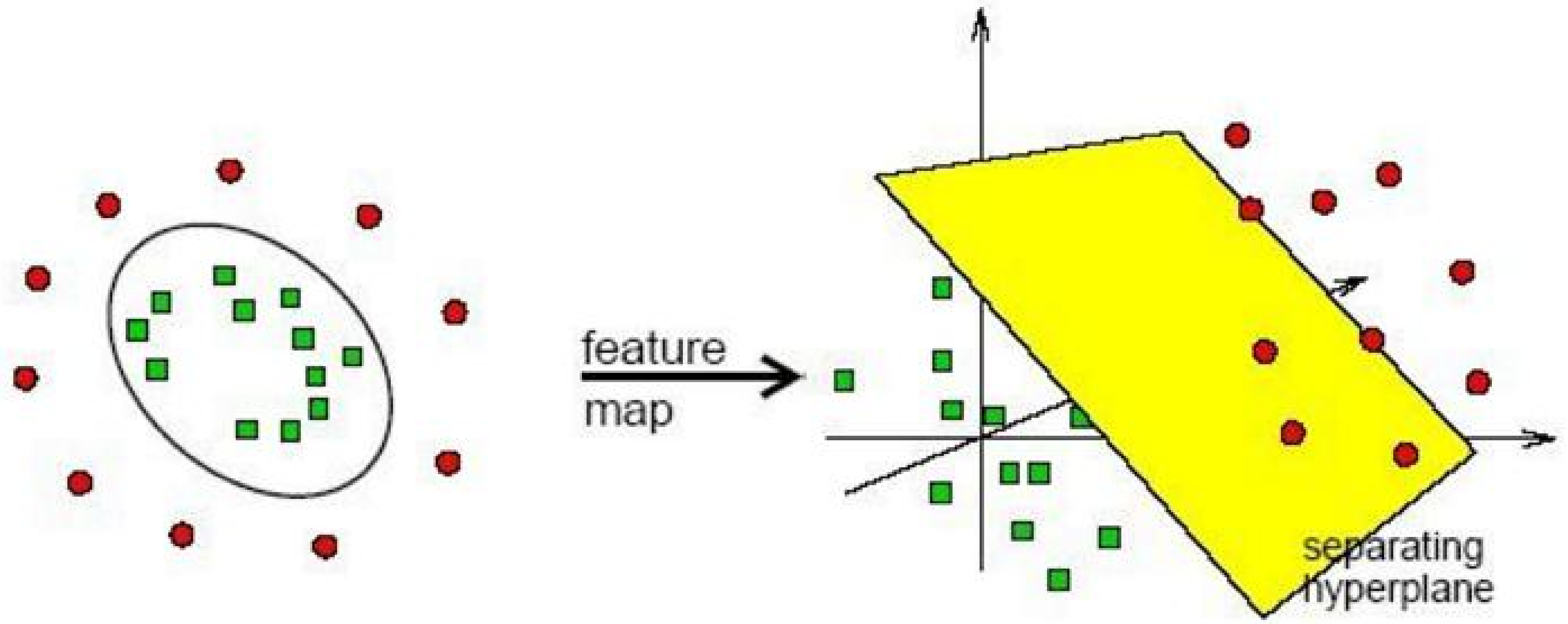
  - 低维（比如我只知道一个人的年龄，性别，那我能对她多了解吗？）
    高维（比如我知道他从出生开始，做过哪些事，赚过哪些钱等)

  - 如果我们对数据更好的了解（是机器去了解他们，我们不需要认识啦)得到的结果不也会更好嘛。

核函数

- 出发点

二维的情况　　　　三维的情况



核函数

- 线性核函数

  - Linear核函数对数据不做任何变换。$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$

  - 何时来使用呢?
    特征已经比较丰富了，样本数据量巨大，需要进行实时得出结果的问题。

  - 不需要设置任何参数，直接就可以用了。

核函数

- 多项式核函数

  - 需要给定3个参数 $(\zeta + \gamma \mathbf{x}^T \mathbf{x}')^Q$ with $\gamma > 0$

  - 一般情况下2次的更常见 $(1 + \gamma \mathbf{x}^T \mathbf{x}')^2$

  - γ(gama)对内积进行放缩，ζ(zeta)控制常数项，Q控制高次项。其特例就是线性核函数了

核函数

还是先从一个小例子来阐述问题。假设我们有俩个数据，x = (x1, x2, x3); y = (y1, y2, y3)，此时在3D空间已经不能对其经行线性划分了，那么我们通过一个函数将数据映射到更高维的空间，比如9维的话，那么 f(x) = (x1x1, x1x2, x1x3, x2x1, x2x2, x2x3, x3x1, x3x2, x3x3)，由于需要计算内积，所以在新的数据在9维空间，需要计算<f(x),f(y)>的内积，需要花费O（n^2）。

在具体点，令x = (1, 2, 3); y = (4, 5, 6)，那么f(x) = (1, 2, 3, 2, 4, 6, 3, 6, 9)，f(y) = (16, 20, 24, 20, 25, 36, 24, 30, 36)，

此时<f(x), f(y)> = 16 + 40 + 72 + 40 + 100+ 180 + 72 + 180 + 324 = 1024

似乎还能计算，但是如果将维数扩大到一个非常大数时候，计算起来可就不是一丁点问题了。

但是发现， K(x, y) = (<x, y>)^2

K(x,y)=(4 + 10 + 18) ^2 = 32^2 = 1024

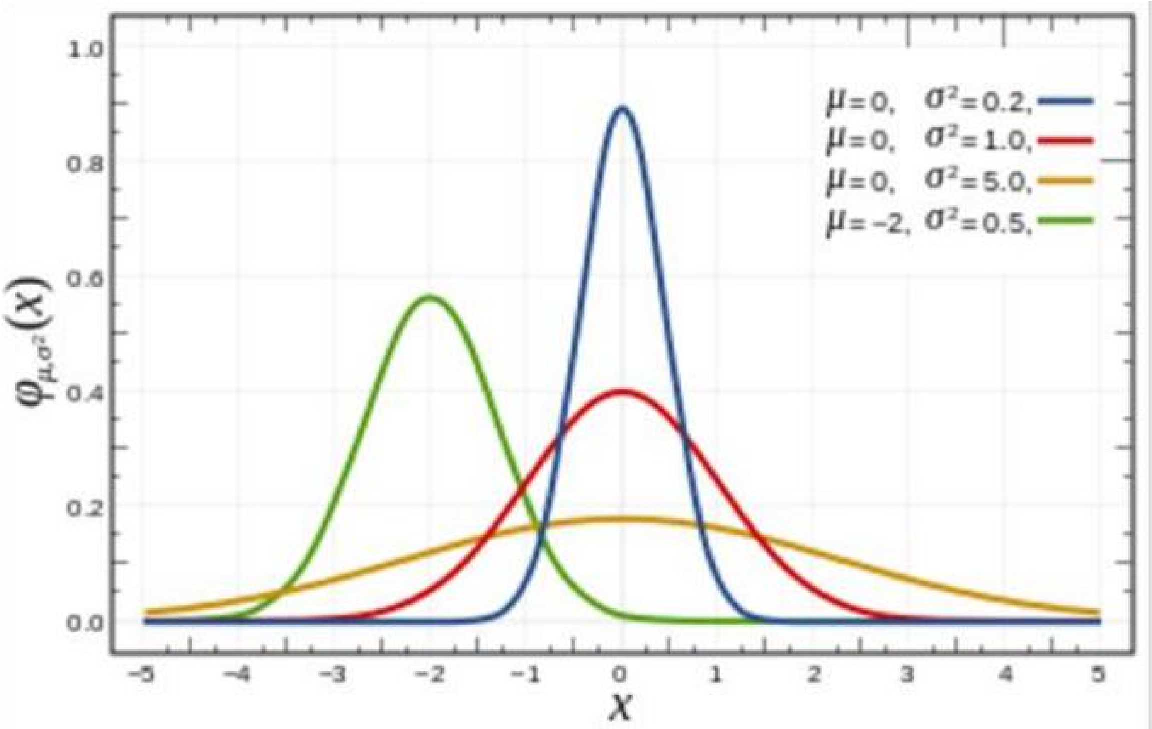
俩者相等，K(x, y) = (<x, y>)^2=<f(x), f(y)>，但是 K(x, y) 计算起来却比<f(x), f(y)>简单的多，也就是说只要用K(x, y)来计算，，效果和<f(x), f(y)>是一样的，但是计算效率却大幅度提高了，如：K(x, y)是O（n），而<f(x),f(y)>是O（n^2）.所以使用核函数的好处就是，可以在一个低维空间去完成高维度（或者无限维度）样本内积的计算，比如K(x,y)=(4 + 10 + 18) ^2 的3D空间对比<f(x), f(y)> = 16 + 40 + 72 + 40 + 100+ 180 + 72 + 180 + 324 的9D空间。

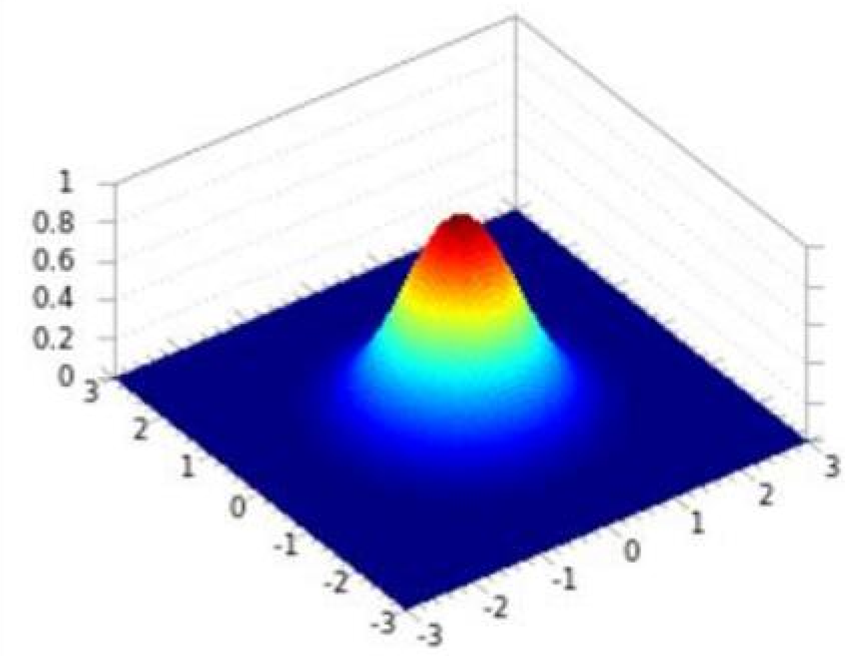
核函数

高斯核函数

一维的高斯 $f(x) = ae^{-\frac{(x-b)^2}{2c^2}}$

二维的高斯 $A\exp\left(-\left(\frac{(x-x_o)^2}{2\sigma_x^2}+\frac{(y-y_o)^2}{2\sigma_y^2}\right)\right)$



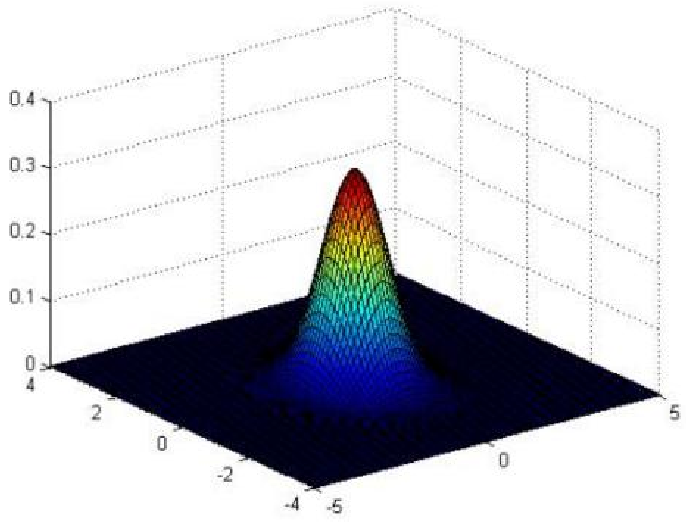


核函数

- 高斯核函数

公式： $K(X,Y) = \exp\left\{-\frac{\|X-Y\|^2}{2\sigma^2}\right\}$

表示什么呢？看起来像是两个样本点之间的距离的度量。
如果X和Y很相似，那结果也就是1了，如果很不相似那就是0了。



核函数

- 高斯核函数

这么做有什么好处呢？能给我做出多少维特征呢？

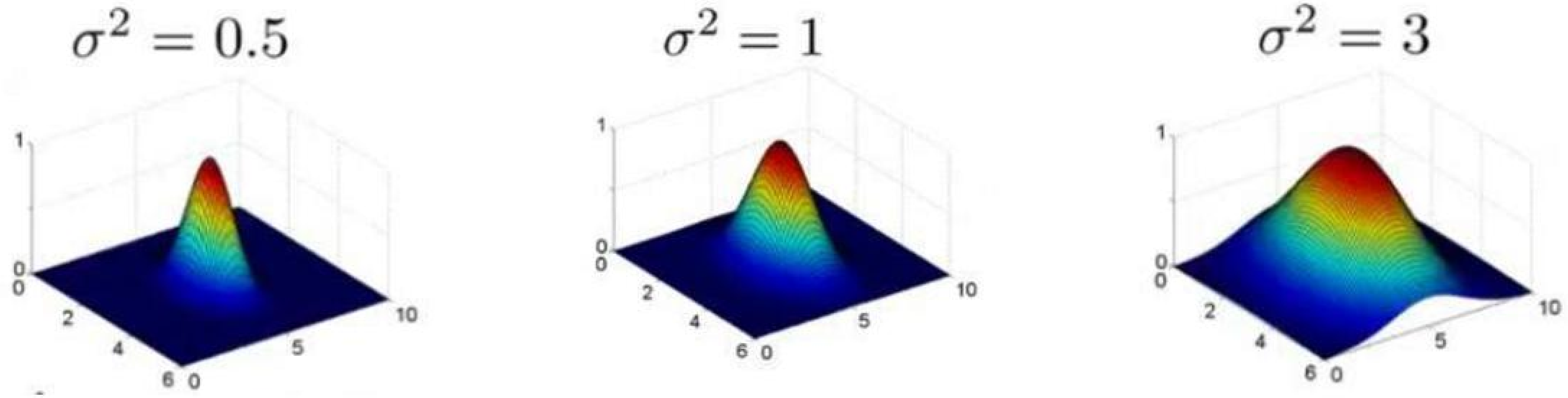
$$
\begin{aligned}
K(x, x') &= \exp(-(x - x')^2) \\
&= \exp(-(x)^2)\exp(-(x')^2)\exp(2xx') \\
&\overset{\text{Taylor}}{=} \exp(-(x)^2)\exp(-(x')^2)\left(\sum_{i=0}^{\infty}\frac{(2xx')^i}{i!}\right) \\
&= \sum_{i=0}^{\infty}\left(\exp(-(x)^2)\exp(-(x')^2)\sqrt{\frac{2^i}{i!}}\sqrt{\frac{2^i}{i!}}(x)^i(x')^i\right) \\
&= \boldsymbol{\Phi}(x)^T\boldsymbol{\Phi}(x')
\end{aligned}
$$

$$
\boldsymbol{\Phi}(x) = \exp(-x^2)\cdot\left(1, \sqrt{\frac{2}{1!}}x, \sqrt{\frac{2^2}{2!}}x^2, \ldots\right)
$$

核函数

- 高斯核函数

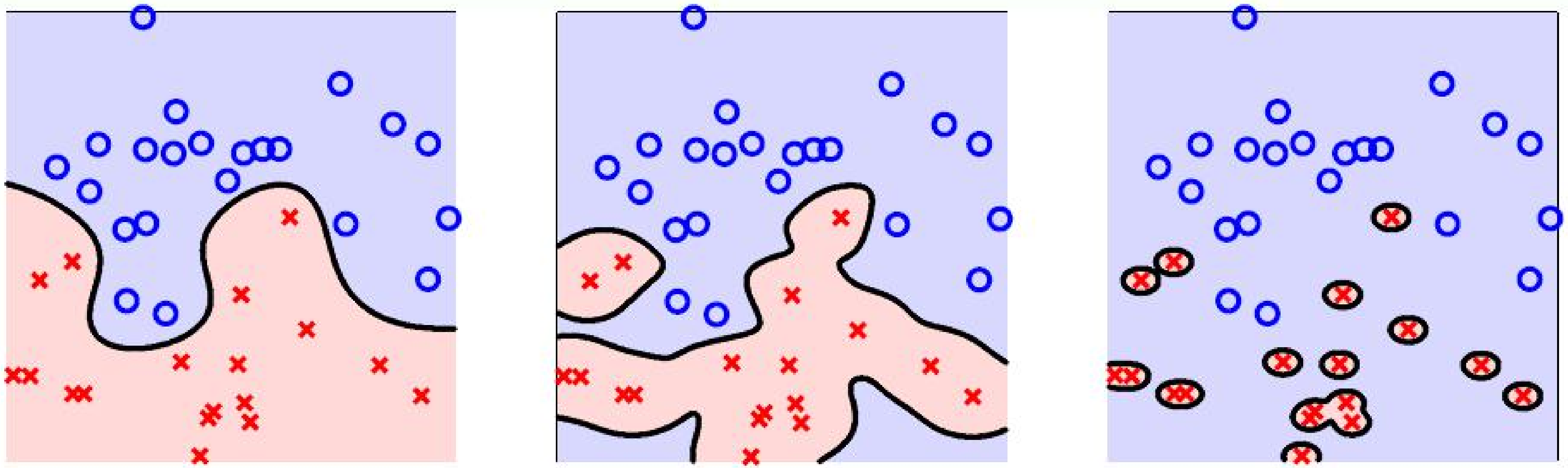
看起来不错，但是它对参数是极其敏感的，效果差异也是很大的！



核函数

- 高斯核函数

决策边界会怎么样呢？（σ越小，切分的越厉害，越容易过拟合）



$\exp(-1\|\mathbf{x}-\mathbf{x}'\|^2)$ $\exp(-10\|\mathbf{x}-\mathbf{x}'\|^2)$ $\exp(-100\|\mathbf{x}-\mathbf{x}'\|^2)$