一、机器学习与相关数学初步

1、Taylor公式及其作用

Taylor公式 – Maclaurin公式

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{f''(x_0)}{2!}(x-x_0)^2+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+R_n(x)$$

$$f(x)=f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\cdots+\frac{f^{(n)}(0)}{n!}x^n+o(x^n)$$

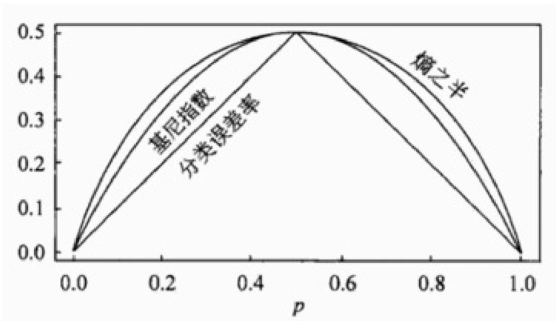
Maclaurin公式是Taylor公式在x=0处进展开行

Taylor公式的作用：泰勒公式把一个复杂的函数可以变成一个简单的多项式相加的形式，可以用来计算某些复杂函数的值。

- 考察基尼指数的图像、熵、分类误差率三者之间的关系
  - 将f(x)=-lnx在x=1处一阶展开，忽略高阶无穷小，得到f(x)≈1-x

$$H(X)=-\sum_{k=1}^{K}p_k\ln p_k$$

$$\approx\sum_{k=1}^{K}p_k(1-p_k)$$

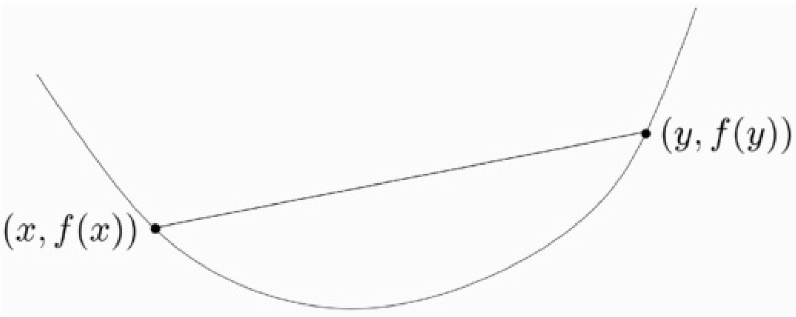


  - 上述结论，在决策树章节中会进一步讨论

2、凸优化

- 若函数f的定义域domf为凸集，且满足

$$\forall x, y \in \operatorname{dom} f, 0 \le \theta \le 1，有$$
$$f(\theta x + (1-\theta) y) \le \theta f(x) + (1-\theta) f(y)$$



凸优化的判定

- 定理：f(x)在区间[a,b]上连续，在(a,b)内二阶可导，那么：
  - 若$f''(x)>0$，则f(x)是凸的；
  - 若$f''(x)<0$，则f(x)是凹的

凸优化的具体应用：

一方面是凸优化的性质，另一方面是凸优化肯定存在一个全局最小值，而且这个利用随机梯度下降不会陷入局部最小值的误区。