如何防止过拟合 +正则化项 （L1L2）

L1各个参数绝对值的和

L2各个参数平方和

**How**

**经验风险：L(W)=1/nΣ(f(x)-y)²**

**结构风险：L(W)=1/nΣ(f(x)-y)²+||w|| （L1）**

**结构风险：L(W)=1/nΣ(f(x)-y)²+||w||² (L2)**

1. ∂L(W)/ ∂Wi = 0 (i=1,2,3…) ①

∂L(W)/ ∂λj = 0 (j=1,2,3…) ②

由①②可得最优解W,λ

1. 若将 W1+W2+..+Wn≤M

或 W1²+W2²+..+Wn²≤M

则限制了W取值范围，即参数受到了约束

1. 所以问题转化为：

**经验风险：minL(W)= 1/nΣ(f(x)-y)²**

S.T. W1+W2+..+Wn≤M

W1²+W2²+..+Wn²≤M

**含不等式约束得最优化问题**

**由KKT求解**

minL(W)=1/nΣ(f(x)-y)²

S.T. ||W||n²-m≤0

**可得拉格朗日函数**

L(W)= 1/nΣ(f(x)-y)²+λ(||W||n²-m)

求偏导

∂L(W)/ ∂Wi = 0 (i=1,2,3…) ①

∂L(W)/ ∂λj = 0 (j=1,2,3…) ②

L2正则与不等书约束最优化问题处在同一个解空间

（W范围相同）

实际上正则化就是将最小化问题转化为带约束的最优化问题！！！