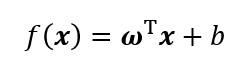
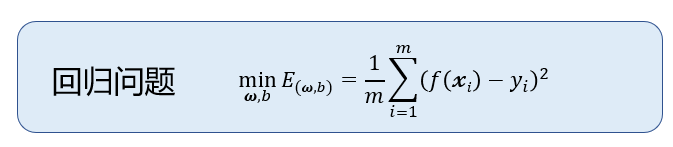
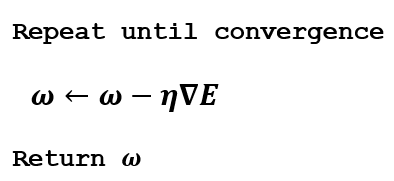
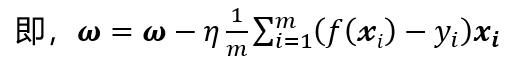


离散属性处理  
>有“序”关系：连续化为连续值  
·{高，矮}：高→1.0；矮→0.0  
>无“序”关系：有k个属性值，则转换为k维向量  
·{男，女}：男→(1,0)，女→(0,1)

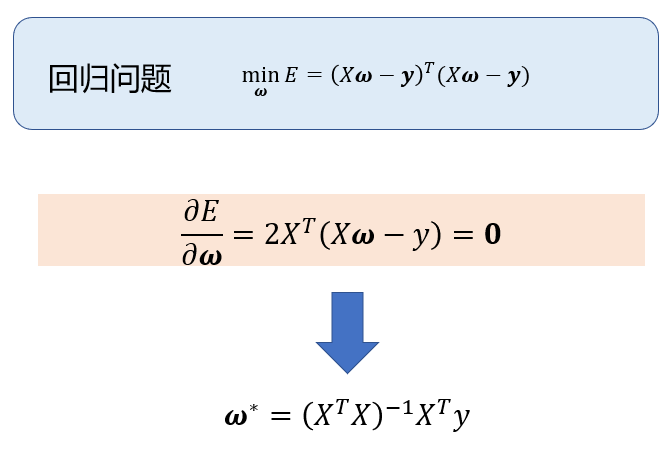
监督学习：有标记。线性模型。

**线性回归**（回归问题：数据标记为连续值）：

1. **最小二乘法**（统计学方法1）：因为E是关于**ω**的凸函数，当它关于**ω**的导数均为零时，得到**ω**的最优解。
2. **梯度下降法**（计算机优化算法）：

随机梯度下降法：每次更新时仅抽取1个样本而非使用全部样本。收敛更快，结果往往在全局最优解附近。

mini-batch梯度下降法：讲所有样本分组，每次更新时抽取1组样本。

1. **最小二乘法的矩阵解法**（统计学方法2）：