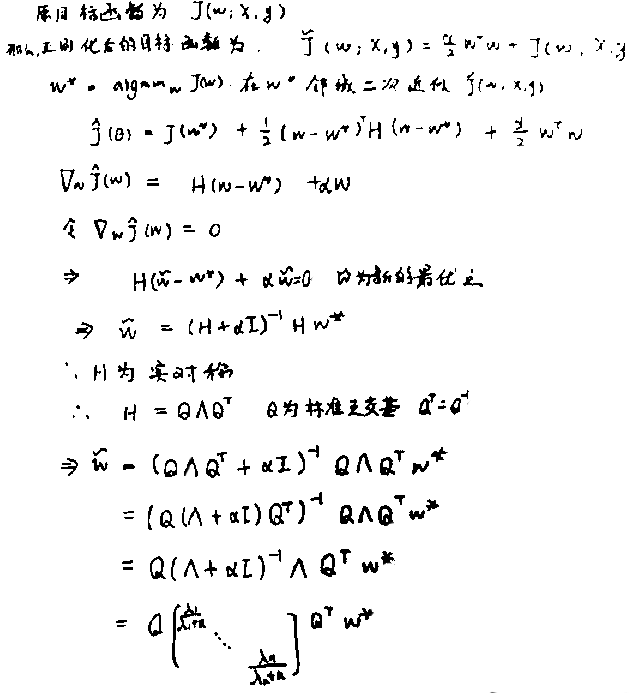
**L1L2范数区别**

//参考《深度学习》7.1

L2范数：

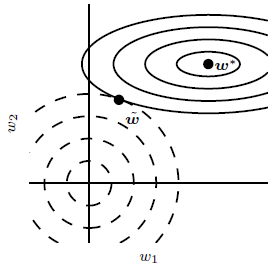


我们可以看到权重衰减的效果是沿着由H 的特征向量所定义的轴缩放 w ∗ 。具体来

说，我们会根据因子缩放与 H 第 i 个特征向量对齐的 w ∗ 的分量。

沿着 H 特征值较大的方向 (如 λ i ≫ α)正则化的影响较小。而 λ i ≪ α 的分量将

会收缩到几乎为零。



经典的用图表示这个推导，这个图中实线椭圆表示没有正则化目标的等值线。虚线圆圈表示L2 正则化项的等值线。实际上是个三维图，自己脑补，能形象的说明2范数正则化在干什么。建议给面试官画三维图。

L1范数：

