第四章 数值计算

4.1 上溢和下溢

下溢（underflow）：当接近零的数被四舍五入为零时发生下溢。

上溢（overflow）：当大量级的数被近似为∞或 $-\infty$时发生上溢。

例子:

问题：softmax函数有上溢和下溢

解决方案：用$softmax(x - max_i x_i)$替代$softmax(x)$可以解决上溢问题和因分母下溢而导致被零除的可能性。

未解决：分子中的下溢仍可以导致整体表达式被计算为零，计算其log时会错误地得到$-\infty$。

解决方案:

$$\log[f(x_i)] = \log\left(\frac{e^{x_i}}{e^{x_1}+e^{x_2}+\cdots e^{x_n}}\right) = \log\left(\frac{\frac{e^{x_i}}{e^M}}{\frac{e^{x_1}}{e^M}+\frac{e^{x_2}}{e^M}+\cdots\frac{e^{x_n}}{e^M}}\right) = \log\left(\frac{e^{(x_i-M)}}{\sum_j^n e^{(x_j-M)}}\right) = \log\left(e^{(x_i-M)}\right)$$
$$-\log\left(\sum_j^n e^{(x_j-M)}\right) = (x_i - M) - \log\left(\sum_j^n e^{(x_j-M)}\right)$$

大家看到，在最后的表达式中，会产生下溢出的因素已经被消除掉了——求和项中，至少有一项的值为1，这使得log后面的值不会下溢出，也就不会发生计算 log(0) 的悲剧。

在很多数值计算的library中，都采用了此类方法来保持数值稳定。

4.2 病态条件

条件数表征函数相对于输入的微小变化而变化的快慢程度。

参考资料：https://blog.csdn.net/u011584941/article/details/44625779

4.3 基于梯度的优化方法

概念:

目标函数、损失函数

一维：导数、梯度下降、临界点(驻点)、局部极小点、局部极大点、鞍点、全局最小点

多维：偏导数、梯度、方向导数、最速下降法(梯度下降)、学习率

4.3.1 梯度之上：Jacobian 和Hessian 矩阵

输入和输出都为向量的函数的所有偏导数组成的矩阵为Jacobian矩阵。

二阶导数组成的矩阵为Hessian矩阵。

Hessian矩阵等价于梯度的Jacobian矩阵。

二阶方向导数的公式：

$$\nabla_u^2 f = \nabla_u(u^T \nabla f) = \nabla_u(\nabla^T f)u = u^T \nabla(\nabla^T f)u = u^T H u$$

参考：https://blog.csdn.net/tina_ttl/article/details/51202566

二阶导数测试

牛顿法

Lipschitz 连续

4.4 约束优化

约束优化、可行点

KKT条件、广义Lagrange函数、等式约束、不等式约束、KKT乘子

4.5 实例：线性最小二乘