无约束最优化：

梯度下降：沿梯度下降 Xn+1 = Xn-Δf’(x)

▽f(X0,Y0)=fx ‘ (x0,y0) i单位向量+f’y(x0,y0)j单位向量

牛顿法：泰勒展开到二阶 Xn+1=Xn – f’’(Xn) / f’(Xn)

共轭法:

拟牛顿法:

约束最优化

最优性条件

可行方向法

惩罚函数法

拉格朗日乘数 : 引入拉格朗日乘子 为了让▽F = λ▽条件 KKT条件

**常见概念**

**目标函数（objective function）**

目标函数就是实现最小化或最大化的目标函数，也被称为评价函数。

**收敛（convergence）**

指经过多步迭代之后得出的数值不应该无限的增大,而是趋于某个数值,

**局部最小值（local mininum）**

局部最小值只是比所有邻近点的函数值都小

**全局最小值（global mininum）**

比定义域内所有其他点的函数值都小，包括了和所有的局部最小值对比

**导数（derivative）**

微积分中重要基础概念，是函数的局部性质，又名微商、导函数值。  
导数在几何中可以求切线，在代数中可求瞬时变化率，在物理中可求速度、加速度。  
一个函数在某一点的导数描述了这个函数在这一点也附近的变化率。 概念如下  
>当函数y=f(x)的自变量x在一点x0上产生一个增量Δx时，函数输出值的增量Δy与自变量增量Δx的比值在Δx趋于0时的极限a如果存在，a即为在x0处的导数，记作f'(x0)或df(x0)/dx。

**一阶导数（first-order derivative）**

描述的是目标函数随输入的变化而变化

**二阶导数（second-order derivative）**

描述的是一阶导数随输入的变化而变化，提供了关于目标函数曲率（curvature）的信息

**梯度（gradient）**

梯度的本意是一个矢量，其方向上的方向导数最大，其大小正好是此最大方向导数.

表示某一函数在该点处的方向导数沿着该方向取得最大值，即函数在该点处沿着该方向（此梯度的方向）变化最快，变化率最大（为该梯度的模）。

梯度的方向是方向导数中取到最大值的方向，梯度的值是方向导数的最大值

曲率影响的是目标函数的下降速度。曲率为正，目标函数会比梯度下降法预期下降得更慢，曲率为负反之。

**无约束优化（unconstrained optimization）**

对自变量的取值没有限制 。例如：梯度下降法（gradient descent）

**约束优化（constrained optimization）**

对x变量取值限制在特定的集合内。例如：线性规划（linear programming）