### 最优化理论

最优化理论是研究函数在给定的一组约束条件下的最小值（或最大值）的数学问题。一般一个最优化问题的标准形式如下：



其中为目标函数，为不等式约束条件，为等式约束条件。

在很多情况下，不等式约束条件可以通过引入新的变量而转化为等式约束条件，因此，最优化问题的一般形式可以简化为仅仅包含等式约束条件的形式：



1）如果目标函数和约束条件都是变量的线性函数，则称该最优化问题为线性规划；

1. 如果目标函数为变量的二次函数，约束条件为变量的线性函数，则称该最优化问题为二次规划；
2. 如果目标函数或约束条件是变量的非线性函数，称该最优化问题为非线性规划。

### KKT条件

条件是指在满足一些有规则的条件下，一个非线性规划问题能有最优解的一个必要和充分条件，这是广义话拉格朗日乘数的成果。所谓最优化条件，是指最优化标准式的最优点必须满足下面的条件：

1. 约束条件满足以及；



1. ，其中为梯度算子；
2. 且不等式约束条件满足：。

上面3个条件的第一个条件是说：最优点必须满足所有等式及不等式限制条件，即

最优点必须是一个可行解；第二、三项表明：在最优点，必须是和的线性组合；和都叫做拉格朗日乘子，不同在于因为不等式限制有方向性，所以每个，而等式约束条件无方向性，所以无符号限制。

### KKT条件的推导（从拉格朗日对偶说起）

我们将优化问题的标准形式称为 ，如下：



其拉格朗日乘子形式为：



定义拉格朗日对偶式为：



很明显，是一组关于的仿射问题，不论原问题是啥形式，对偶问题一定是凸优化问题。（性质优良，例如极值点唯一。）

令：



对于满足约束条件的而言，，所以的最大值：



这样，就将原始带约束的优化问题转化为如下无约束问题；对原问题，写成如下形式：



什么是对偶问题 ？

形式上只是将上式的与交换位置：

 ：



根据拉格朗日对偶的定义有：



则，对偶问题变为：在可行域范围内寻找最优的使得最大，即：



交换之后， 与 并不相等，但为什么要这么做？

其实就是因为原问题是关于的带约束优化问题，如果不是二次凸优化的话（往往现实中就是如此）比较难解，复杂麻烦，而对偶问题则简单很多，就是个线性凸优化问题。实际上，无论原问题是什么形式，对偶问题总是个凸优化的问题，其极值是唯一的。

有一个很好的性质：它是 的一个下界，即若，那么对于所有的和，有：。

证明如下：

对于满足约束条件的，总有：



则有：



即对于最优点有：。

实际上，我们确定了的下界性质：，对于对偶问题：



的求解也是在逼近原问题的最优解。

**1）**若 的最优解为，则有：。

这个性质就是弱对偶，对于所有的优化问题都是成立的，其中为 。

弱对偶有什么作用？

对于一些很难求解的原始问题（甚至是问题），我们可以通过找出其对偶问题，通过优化这个对偶问题得到原始问题的一个下界估计，或者有时候都不用优化这个对偶问题，而是通过一些方法（随机法）选取一些和，代入，也会得到一些下界，只是不一定是最大的那个。

**2）**强对偶和限制条件。

当时，就是强对偶成立的条件，此时，我们可以通过求对偶问题的最优解找到原始问题的最优解。

一般而言，强对偶是很难的，但是如果原始问题是凸问题且满足条件的话，强对偶就成立（这只是强对偶成立的一种情况，但不是唯一情况，某些非凸问题强对偶也成立）。

条件：指存在严格满足约束条件的点，（严格是指：），即：



假设为原始问题的极值点，是对偶问题的极值点，且强对偶成立：，则：



因为上式两头相等，所以可将不等号换成等号：



可知：是的一个极值点，则在处的梯度等于零：



由第二个不等号可知，是的，等式成立时，有：



上面这个称为：互补松弛条件。

将上述条件写到一起就是条件：



条件的作用在于：某一类问题的条件可以直接被解出，而对于大多数其他优化问题，就把任务由解优化问题转移到解条件上，因为二者等价。任何满足强对偶的问题都满足条件，当原始问题是凸优化问题，且存在和满足条件，则他们分别是原始问题和对偶问题的极值点，并且强对偶成立。

总的来说，一个优化问题，通过求出它的对偶问题，在只有弱对偶成立的情况下，我们至少可以得到一个原始问题的下界，而如果强对偶成立，则可以直接通过求解对偶问题来解决原始问题。



2017.07.28.night