这周主要学习了优化理论中的坐标下降法（Coordinate Descent），交替方向乘子法（Alternating Direction Method of Multipliers）。

1：坐标下降(CD)法

坐标下降是一个非梯度优化算法。与梯度优化算法沿着梯度最速下降的方向寻找函数最小值不同，坐标下降法依次沿着坐标轴的方向最小化目标函数值。坐标下降的和核心思想是将优化问题分解，通过迭代将大多数的自变量固定，针对剩余的自变量求极值。这样一个高维的优化问题就分解为了多个一维的优化问题。

若 f（x）是可微的凸函数，那么局部极小值就是全局最小值。

假设我们有目标函数 f(x,y)=5x^2-6xy+5y^2 需要求最小值，起始点为(-0.5，-1.0），此时函数值为 3.25。

此时固定 x ,可得 f(y|x=-0.5)=5y^2+3y+1.25 ,再通过求导求极值可得 y=-0.3。所以可得新的点为（-0.5，-0.3），f=0.8；

然后固定 y，对于 f（y=-0.5|x）进行求导，求的 x 的值，经过多次迭代可得最后的值。

2：交替方向乘子法(ADMM)

ADMM 通常用于解决存在两个优化变量的只含等式约束的优化类问题

一般形式为：

$$\min_{x,z} \quad f(x) + g(z)$$
$$s.t. \quad Ax + Bz = c$$

为了解决此类凸优化问题，需要定义增广拉格朗日函数。

$$L_\rho(x,z,u) = f(x) + g(z) + u^T(Ax + Bz - c) + \frac{\rho}{2}||Ax + Bz - c||_2^2$$

然后每一步只更新一个变量而固定另外两个变量，如此交替重复更新。

$$
\begin{aligned}
&step1: \quad x^{(k)} = arg\min_x L_\rho(x, z^{(k-1)}, u^{(k-1)}) \\
&step2: \quad z^{(k)} = arg\min_z L_\rho(x^{(k)}, z, u^{(k-1)}) \\
&step3: \quad u^{(k)} = u^{(k-1)} + \rho(Ax^{(k)} + Bz^{(k)} - c)
\end{aligned}
$$

然后根据终止条件来停止迭代。