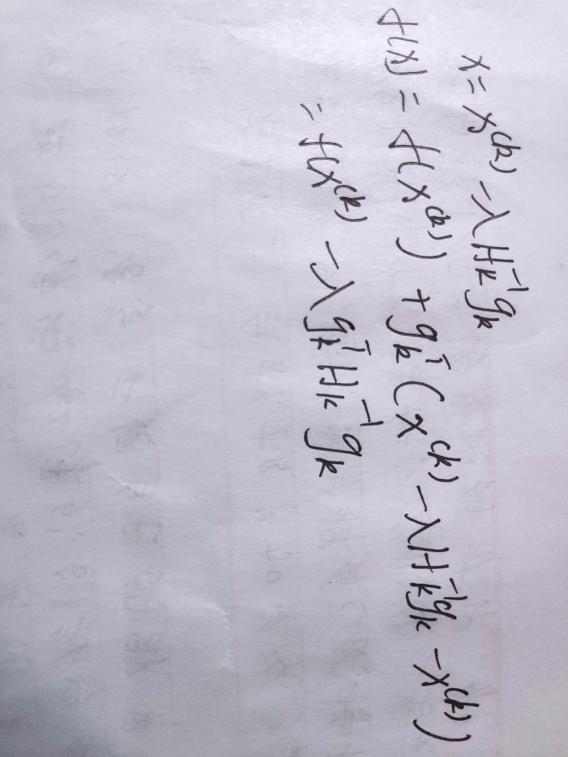
1. **自己提出的问题的理解：**

**问题1：**如何从B.14推到B.15的？

**讨论后的理解：**



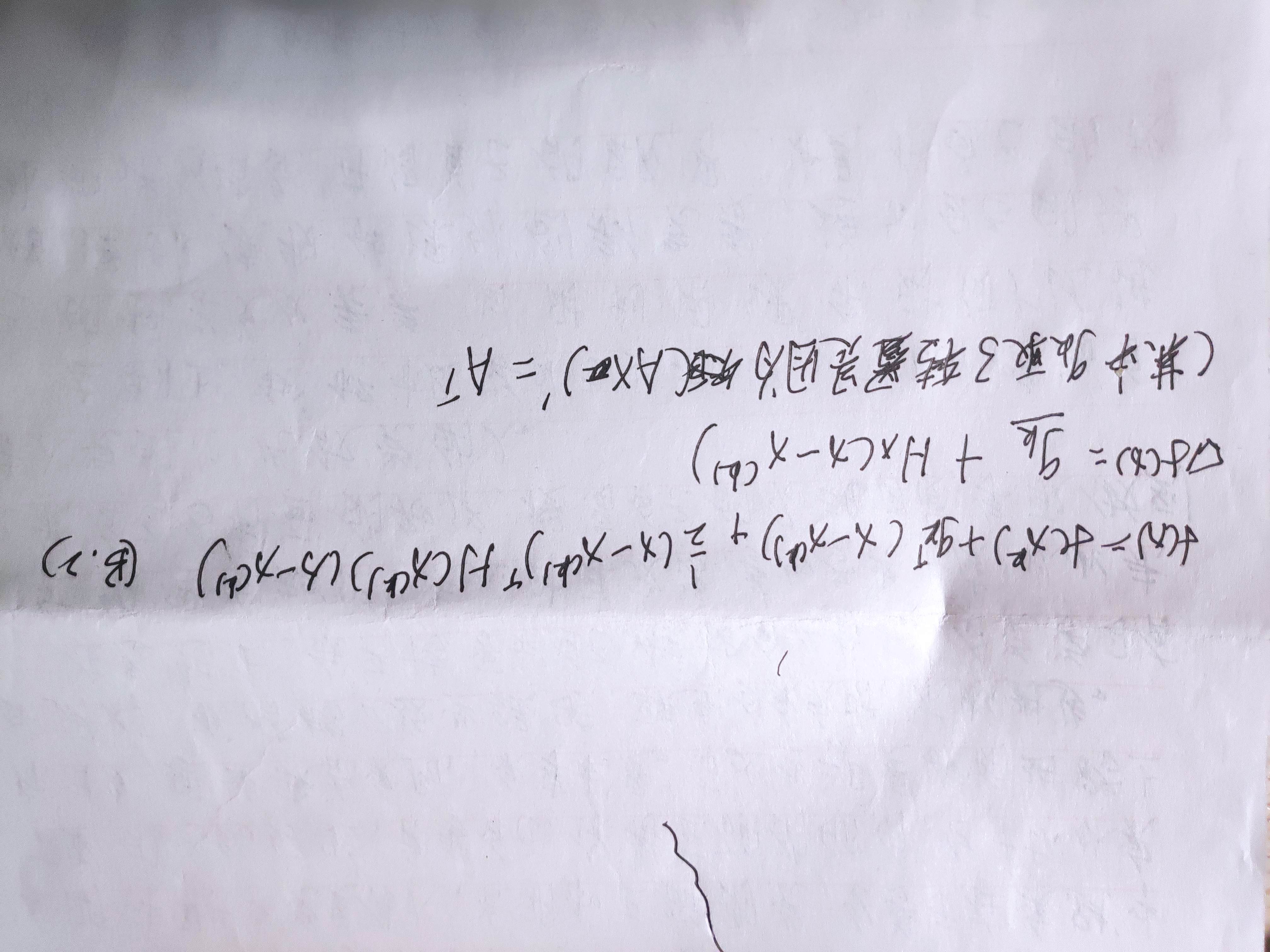
**问题2：**拟牛顿法中近似矩阵的初始值G0和B0需要预先设定好嘛？还是说可以通过什么方式求出来？

**讨论后的理解：**首先，Gk和Bk在迭代的过程中都需要是正定矩阵，其次，如书中所说，如果初始矩阵G0是正定的，则迭代过程中的每个矩阵Gk都是正定的。B同理，所以对于G0和B0来说，只要随便取一个正定矩阵，就可以满足Gk是正定矩阵的条件。

1. **别人提出的问题的理解：**

**问题1：**442，式B.2是怎么推式B.6的

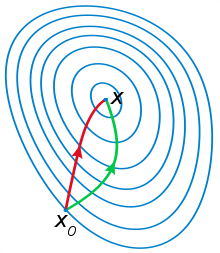
**讨论后的理解：**



参考内容：https://blog.csdn.net/lisa\_ren\_123/article/details/81983785

**问题2：**从数学上来说，为什么牛顿法和拟牛顿法能够更快的找到极小值

**讨论后的理解：**因为牛顿法和拟牛顿法是二阶收敛，而梯度下降法是一阶收敛，简单来说，和求极小值等操作相同，一阶收敛在求梯度的时候会在当前位置选择一个最大的坡度去走，而二阶收敛则不仅能够选择一个最大的坡度去走，同时也会尽可能的保证下次也能够选择到全局最大坡度。



如图所示，其中绿色的路径代表梯度下降法，红色的路径代表牛顿法和拟牛顿法。

**问题3：**为什么凸函数时，梯度下降法是全局最优解？

**讨论后的理解：**因为凸函数只有一个最低谷，和那种“道路曲折”的曲线不同，凸函数只要能够一直往梯度最大的方向去移动，一定能够移动到最低点，同时最低点的时候梯度为0，所以会收敛到最优解。

1. **读书计划**

1、本周完成的内容章节：《统计学习方法》附录ABC

2、下周计划：《统计学习方法》第7章

1. **读书摘要及理解或伪代码的具体实现**
2. **读书摘要：**

1.原始问题和对偶问题会在同时取得最优解。

2.在单变量的函数中，梯度其实就是函数的微分，代表着函数在某个给定点的切线的斜率

在多变量函数中，梯度是一个向量，向量有方向，梯度的方向就指出了函数在给定点的上升最快的方向

3.如果需要往反方向优化的时候，只需要在梯度前面加一个负号就好了。

1. **代码实现**

def func\_1d(x):

    return x \*\* 2 + 1

def grad\_1d(x):

    return x \* 2

def gradient\_descent\_1d(grad, cur\_x, learning\_rate, precision,max\_iters):

    for i in range(max\_iters):

        grad\_cur = grad(cur\_x)

        if abs(grad\_cur) < precision:

            break

        cur\_x = cur\_x - grad\_cur \* learning\_rate

        print("第", i, "次迭代：x 值为 ", cur\_x)

    return cur\_x

if \_\_name\_\_ == '\_\_main\_\_':

    gradient\_descent\_1d(grad\_1d, cur\_x=10, learning\_rate=0.2, precision=0.0001, max\_iters=100)  
Output

第 0 次迭代：x 值为 6.0

第 1 次迭代：x 值为 3.5999999999999996

第 2 次迭代：x 值为 2.1599999999999997

第 3 次迭代：x 值为 1.2959999999999998

第 4 次迭代：x 值为 0.7775999999999998

第 5 次迭代：x 值为 0.46655999999999986

第 6 次迭代：x 值为 0.2799359999999999

第 7 次迭代：x 值为 0.16796159999999993

第 8 次迭代：x 值为 0.10077695999999996

第 9 次迭代：x 值为 0.06046617599999997

第 10 次迭代：x 值为 0.036279705599999976

第 11 次迭代：x 值为 0.021767823359999987

第 12 次迭代：x 值为 0.013060694015999992

第 13 次迭代：x 值为 0.007836416409599995

第 14 次迭代：x 值为 0.004701849845759997

第 15 次迭代：x 值为 0.002821109907455998

第 16 次迭代：x 值为 0.0016926659444735988

第 17 次迭代：x 值为 0.0010155995666841593

第 18 次迭代：x 值为 0.0006093597400104956

第 19 次迭代：x 值为 0.0003656158440062973

第 20 次迭代：x 值为 0.0002193695064037784

第 21 次迭代：x 值为 0.00013162170384226703

第 22 次迭代：x 值为 7.897302230536021e-05

第 23 次迭代：x 值为 4.7383813383216124e-05