TLS：

针对有误差的数据矩阵和数据向量A,b需满足：

C:\Users\Lenovo\AppData\Local\Temp\ksohtml43196\wps1.jpg,

C:\Users\Lenovo\AppData\Local\Temp\ksohtml43196\wps7.jpg

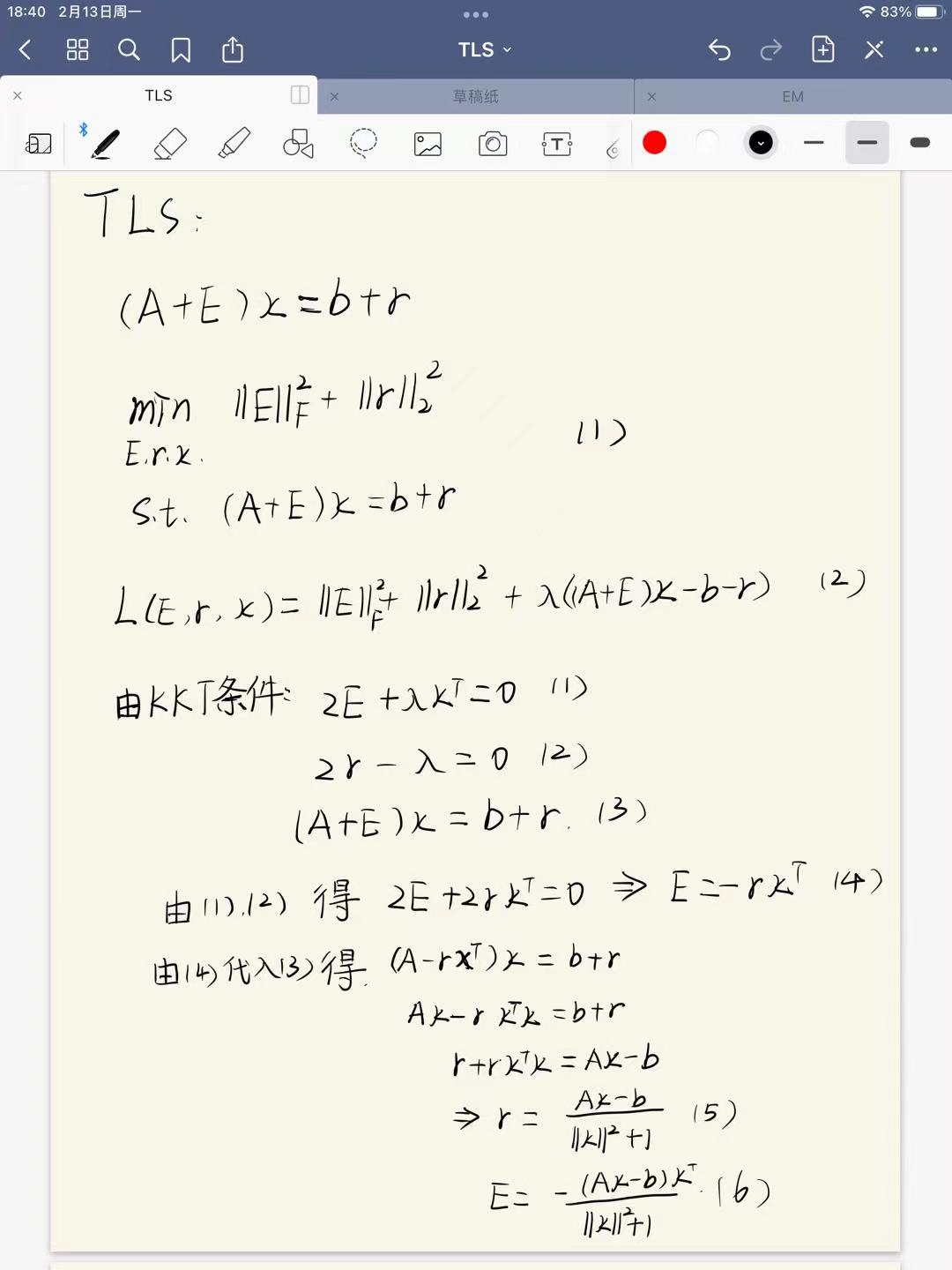
构建拉格朗日函数：

C:\Users\Lenovo\AppData\Local\Temp\ksohtml43196\wps8.jpg, C:\Users\Lenovo\AppData\Local\Temp\ksohtml43196\wps9.jpg

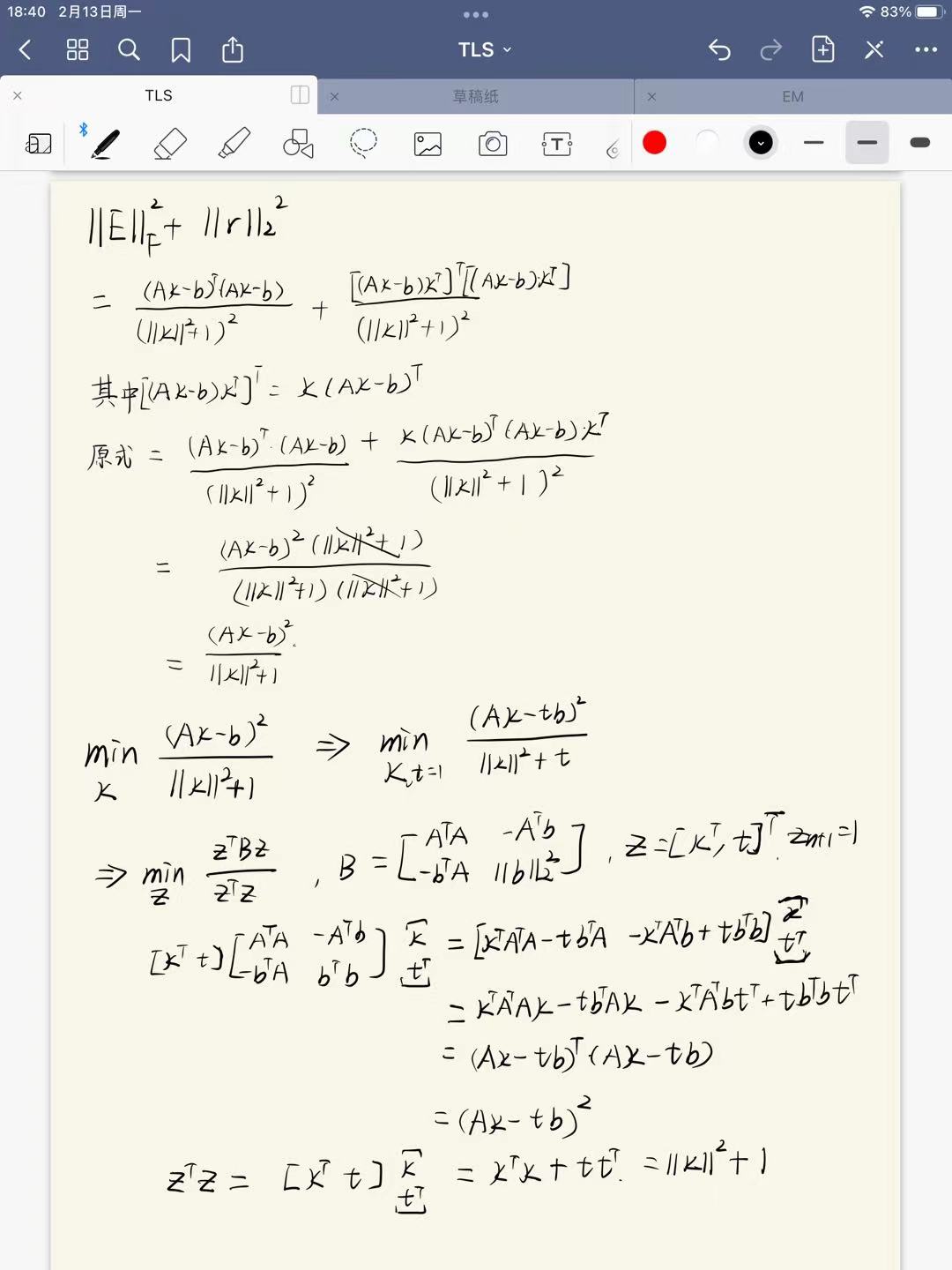
KKT Condition: C:\Users\Lenovo\AppData\Local\Temp\ksohtml43196\wps10.jpg (1), C:\Users\Lenovo\AppData\Local\Temp\ksohtml43196\wps11.jpg (2), C:\Users\Lenovo\AppData\Local\Temp\ksohtml43196\wps12.jpg (3)

根据(1)和(2):  (4),

将(4)代入(3)得到： (5),  (6)



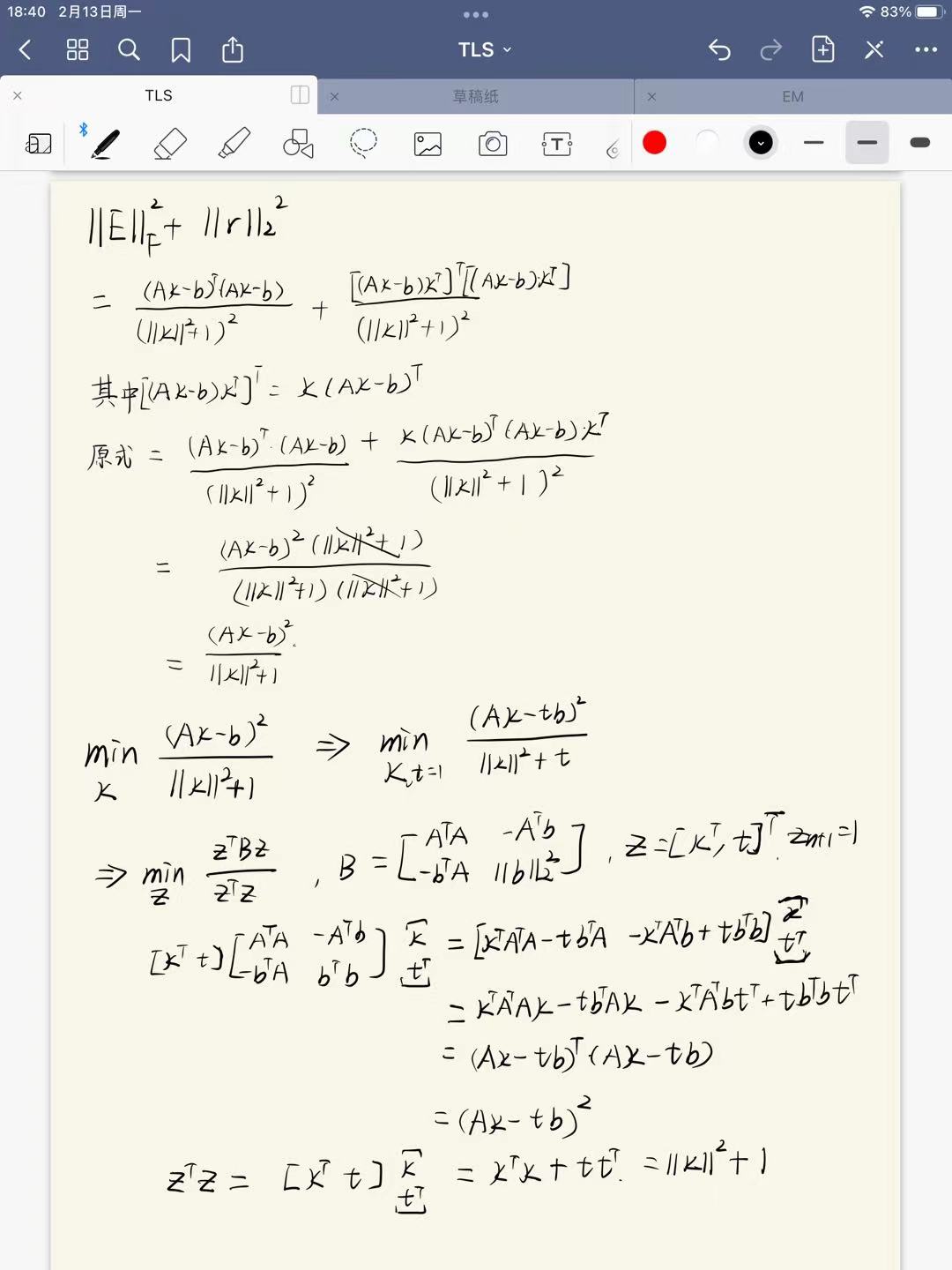
那么目标函数就可以转化为：



首先，利用homogenization argument将问题等价转化为：

此时，再令C:\Users\Lenovo\AppData\Local\Temp\ksohtml43196\wps14.jpg

原问题转化成为



令，构建拉格朗日函数：

根据KKT条件，我们解得： 

可以看出我们所求得最优解z是B的最小特征值对应的特征向量

而根据可以求得：

**KKT条件**：

拉格朗日函数：



其中：（1）表示在取值最优的点f(x)和约束区域边界函数的梯度大小相等，方向相反

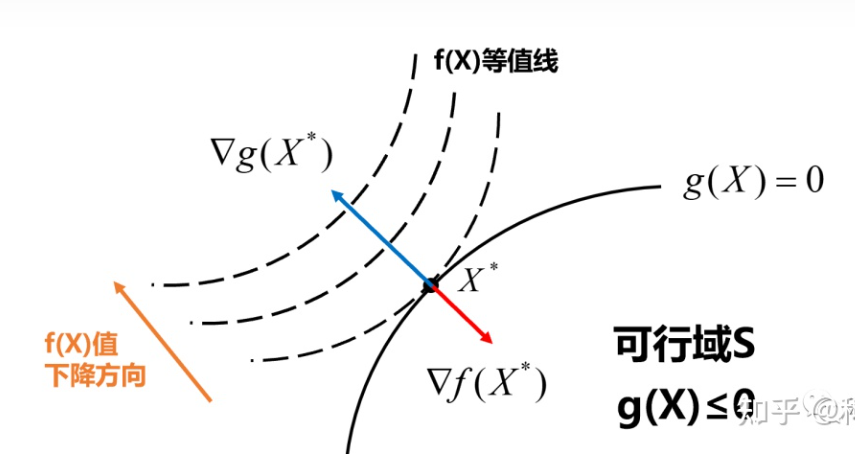
（2）表示在最优值时，起到约束作用的的函数满足g(X)=0,没有约束作用的对应的λ为0.两者不能同时为0.

（3）满足原始等式约束

（4）见图示

（5）满足原始不等式约束

（等式求解，不等式排除解）



**MLE推导线性回归：**

最小二乘估计：

但数据都带有一定噪声，假设噪声满足零均值的高斯分布

因此，

则

使用MLE求解W：

首先是最大似然函数：



再求W：



投硬币：

1. 设置初始值：，
2. 根据，和观察的结果分别计算出硬币来自A、B的概率并求出期望
3. 根据期望更新，直至收敛

EM：

给定数据集，假设样本间相互独立，我们想要拟合模型到数据的参数。根据分布得到似然函数：

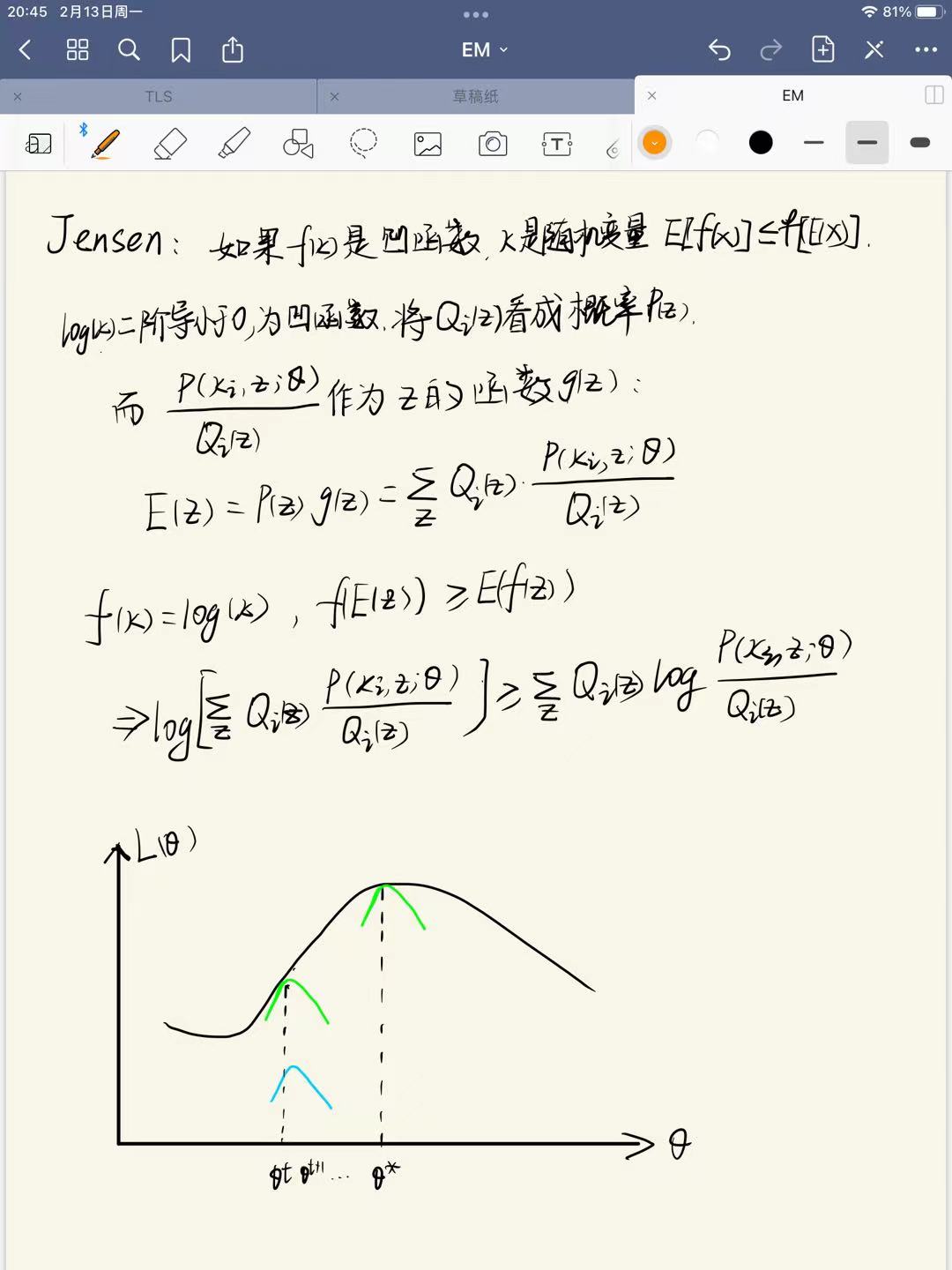
我们可以最大化似然函数以求解，但当存在隐变量时，似然函数表示为：



对于某个样本i，我们使用表示样本i隐含变量z的分布概率，则

，且

引入，并利用Jensen不等式：

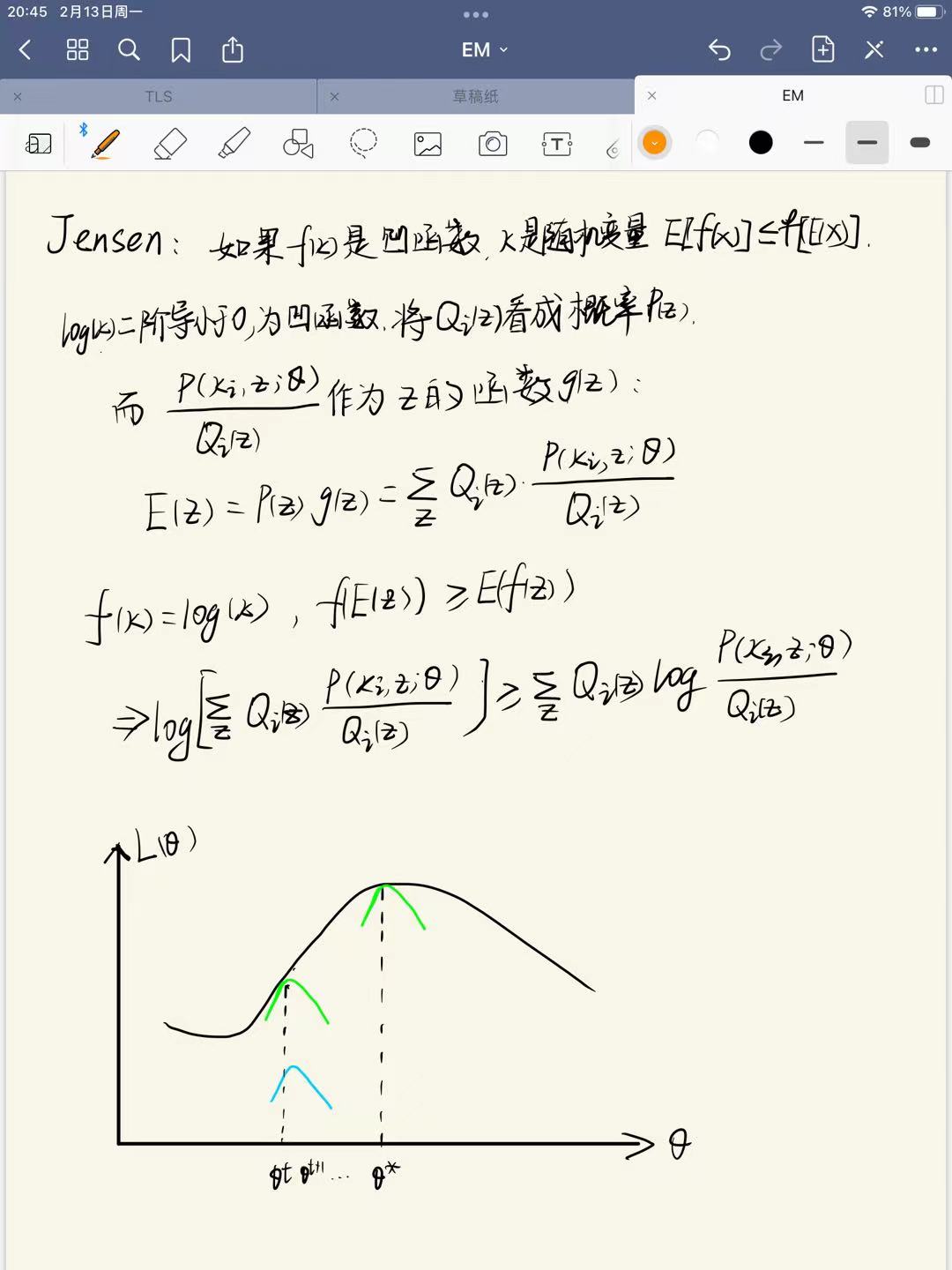




似然函数变换为：



我们得到了的形式，如图所示：



纵向变化：固定改变的过程

横向变化：固定改变的过程

不等式取等号：当为常数时，不等式可以取等号

，对等式变换并两边求和





E-Step：求条件期望，本质是求出来自哪种类别的概率，为简化将log中的去掉，

关于Q(z)：



两边同时乘以Q(z)且求和即可得到期望：



可以发现：

M-Step: 求，更新参数

K-means:



初始化

更新，

对参数求导等于0，得

得到新一轮的参数

再根据此更新

GMM：

对数似然函数：，表示样本属于第k个高斯函数的概率，且满足

E:求期望，本质是求出，即

M: