Homework 10

1、最小二乘的目标是$\min\|Ax-b\|_2$，证明最优解x^*满足$A^TAx=A^Tb$

$\|Ax-b\|^2$对x求导，$A^T(Ax-b)=0\Rightarrow A^TAx=A^Tb$

2、在课上，我们介绍了一种求近似最小二乘的方法：随机选择一个矩阵S，计算SA和Sb，然后求解$\min_{x\in R^d}\|SAx-Sb\|_2$得到$x'$作为精确解$x^*$的近似解。我们证明了如果$S$是$(d+1)$-维空间$[A,b]$的子空间嵌入（Subspace Embedding），那么x'有$\geq 1-\delta$的概率满足$\|Ax'-b\|_2\leq(1+\epsilon)\min_{x\in R^d}\|Ax-b\|_2$。之前我们要求$S$至少要有$d/\epsilon^2$行，现在我们将说明$S$可以只要$O(d/\epsilon)$行。

(a) 令$U\in R^{n\times r}$为A矩阵列向量张成空间的标准正交基（orthonormal basis），其中$r=\text{rank}(A)$。证明如果$x'=\arg\min_{x\in R^r}\|SUx-Sb\|_2$满足$\|Ux'-b\|_2\leq(1+\epsilon)\min_x\|Ux-b\|_2$，那么$y'=\arg\min_{y\in R^d}\|SAy-Sb\|_2$满足$\|Ay'-b\|_2\leq(1+\epsilon)\min_y\|Ay-b\|_2$

因为U是A矩阵列向量张成空间的标准正交基，所以我们可以把Ay'写成$Ay'=Ux$，$x\in\mathbb{R}^r$。因为$x'=\arg\min_{x\in R^r}\|SUx-Sb\|_2$，所以$\|SUx'-Sb\|_2\leq\|SAy'-Sb\|_2$。

同样因为U和A列向量的张成空间相同，所以我们可以把Ux'写成$Ux'=Ay$，$y\in\mathbb{R}^d$。因为$y'=\arg\min_{x\in R^d}\|SAy-Sb\|_2$，所以$\|SAy'-Sb\|_2\leq\|SUx'-Sb\|_2$。所以$\|SUx'-Sb\|_2=\|SAy'-Sb\|_2$。$\|Ay'-b\|_2=\|Ux'-b\|_2$($S$是子空间嵌入)

类似的，$\min_x\|Ux-b\|_2=\min_y\|Ay-b\|_2$。

于是有$\|Ay'-b\|_2=\|Ux'-b\|_2\leq(1+\epsilon)\min_x\|Ux-b\|_2=(1+\epsilon)\min_y\|Ay-b\|_2$

(b) 因此我们只要证明对于$O(d/\epsilon)$行的矩阵S，$x'=\arg\min_{x\in R^d}\|SUx-Sb\|_2$能够推出$\|Ux'-b\|_2\leq(1+\epsilon)\min_x\|Ux-b\|_2$。令$x^*=\arg\min_x\|Ux-b\|_2$,证明$\|Ux'-b\|_2^2=\|Ux^*-b\|_2^2+\|U(x'-x^*)\|_2^2$

根据勾股定理有$\|Ux'-b\|_2^2=\|UU^Tb-b\|_2^2+\|Ux'-UU^Tb\|_2^2$，意思是$b$到$Ux$的距离等于$b$到$U$投影的距离+投影到$U'x$的距离。

因为$x^*=\arg\min_x\|Ux-b\|_2\Rightarrow x^*=U^Tb$，代入得证。

(c)* 证明如果S是一个由独立同分布 (i.i.d., independent and identical distribution)，均值为0，方差为$1/k$的高斯随机变量组成的$k\times n$的矩阵，其中$k=O(d/\epsilon)$，那么$\|U(x'-x^*)\|_2^2=O(\epsilon)\|Ux^*-b\|_2^2$

你在证明中会需要用到：(1) S是一个任意确定的d-维子空间的$(1\pm 1/2)$近似子空间嵌入；(2) S满足近似矩阵乘（approximate matrix product），这两项可以作为结论直接使用。

你可能还需要用到x和x'的特定形式：$x'=(SU)^-Sb$，$x^*=U^Tb$，对于一个列线性独立的矩阵C，$C^-=(C^TC)^{-1}C^T$

因为S是U的$O(1)$-近似子空嵌入，且列互相独立，所以SU列互相独立。于是$(SU)^-=((SU)^TSU)^{-1}(SU)^T=(U^TS^TSU)^{-1}U^TS^T$，$x'=(U^TS^TSU)^{-1}U^TS^TSb$，$x^*=U^Tb$。

$$\begin{aligned}\|U(x'-x^*)\|_2^2&=O(1)\|U(U^TS^TSU)^{-1}U^TS^TSb-UU^Tb\|_2^2\\&=O(1)\|(U^TS^TSU)^{-1}U^TS^TSb-U^Tb\|_2^2\end{aligned}$$

因为S是一个任意确定的d-维子空间的$(1\pm 1/2)$近似子空间嵌入，所以$(U^TS^TSU)^{-1}$的元素取值范围都属于$[2/3,2]$，于是

$$
\begin{aligned}
\|(U^TS^TSU)^{-1}U^TS^TSb - U^Tb\|_2^2 &= O(1)\|(U^TS^TSU)((U^TS^TSU)^{-1}U^TS^TSb - U^Tb)\|_2^2 \\
&= O(1)\|U^TS^TSb - U^TS^TSUU^Tb\|_2^2 \\
&= O(1)\|U^TS^TS(b - Ux^*)\|_2^2
\end{aligned}
$$

现在使用近似矩阵乘

$$
\|U^TS^TS(b - Ux^*)\|_2^2 = O(\epsilon/d)\|U^T\|_F^2 \cdot \|Ux^* - b\|_2^2 = O(\epsilon)\|Ux^* - b\|_2^2
$$