1. 一个广泛用来估计一般矩阵A的特征值的方法是QR算法。在适当的条件下，该算法产生一个矩阵序列，序列中的矩阵全部相似于A，矩阵几乎是上三角的，并且主对角线上的元素近似于A的特征值。将A分解为$A=Q_1R_1$,其中$Q_1^T=Q_1^{-1}$，而R_1是上三角矩阵。将Q_1与R_1交换形成$A_1=R_1Q_1$，继续分解为 $A_1=Q_2R_2$, 令$A_2=R_2Q_2$，一直做下去。

- 证明(10分)：A相似于$A_1, A_2, \ldots$。（10分）
- 验证(20分)以A为例，验证该算法。

$$A=\begin{pmatrix} 4 & -3 & 2 \\ 1 & -2 & 3 \\ -1 & 1 & 0 \end{pmatrix}$$

- (10分) Python中 $numpy.roots(p)$是用来计算多项式方程$p(t)=0$的根的。阅读说明文档，阐述roots命令所用到的算法的基本思想。

2.证明(20分)：$x=A^+b$是集合$\arg\min_x ||Ax-b||$中$||x||_2$最小的。

3.证明(20分)：给定$X,Y\in R^{n\times p}$是同一批数据在不同角度下的测量结果,寻找正交阵$Q\in R^{p\times p}$使得

$$||X-YQ||$$

最小。证明如果 $Y^TX=U\Sigma V^T$, 那么 $Q=UV^T$。

4.求解(20分)：设$Q(x)=7x_1^2+x_2^2+7x_3^2-8x_1x_2-4x_1x_3-8x_2x_3$，在条件$x_1^2+x_2^2+x_3^2=1$的限制下，求出向量$x$，是的$Q(x)$最大。

1 解：

1) 证明：$A = Q_1R_1 \ \Rightarrow\ Q_1^{-1}AQ_1 = R_1Q_1 = A_1$, 故 A 相似于 A_1.
同理，$A_1 = Q_2R_2 \ \Rightarrow\ Q_2^{-1}A_1Q_2 = Q_2^{-1}Q_1^{-1}AQ_1Q_2 = R_2Q_2 = A_2$, 故 A 相似于 A_2.
同理可证，得 A 相似于 $A_1, A_2, \ldots$。

2) $(A-\lambda I)\alpha = 0 \Rightarrow |A-\lambda I| = 0 \Rightarrow \begin{vmatrix} 4-\lambda & -3 & 2 \\ 1 & -2-\lambda & 3 \\ -1 & 1 & -\lambda \end{vmatrix} = 0 \Rightarrow \lambda = 3.3465, 0.7222, -2.0687$

对 A 进行 QR 算法，$A = Q_1R_1$，其中

$$Q_1 = \begin{pmatrix} -0.9428 & 0.2722 & 0.1925 \\ -0.2357 & -0.9526 & 0.1925 \\ 0.2357 & 0.1361 & 0.9623 \end{pmatrix},\ R_1 = \begin{pmatrix} -4.2426 & 3.5355 & -2.5927 \\ 0 & 1.2247 & -2.3134 \\ 0 & 0 & 0.9623 \end{pmatrix}$$

令 $A_1 = R_1Q_1$，再进行 QR 分解得 $A_1 = Q_2R_2$，重复上述过程，迭代 10 次得

$$A_{10} = \begin{pmatrix} 3.3549 & -3.4007 & 2.9091 \\ 0.0134 & -2.0771 & 2.9801 \\ 0 & 0 & 0.7222 \end{pmatrix}$$

其主对角线元素近似于 A 的特征值，算法验证毕.

3) 对于 numpy.roots(p) 函数，用来计算多项式方程的根，如果输入的是标量，则将其转换为数组形式，找到所有的非零数组项，对于其他尾零个数即为根为 0 的个数，去掉前导和尾零项后转化为浮点数进行计算，对多项式方程构造友矩阵，通过对友矩阵特征值的求解，得到的即是原方程的解.

2 证明：设最小二范数解为 x，即有 $min_x||x||_2^2 = x^Tx\ (s.t. Ax=b)$. 引入拉格朗日算子 $J(x) = \frac{1}{2}x^Tx - \lambda(Ax-b)$，对该式求导 $\nabla J(x) = x - A^T\lambda = 0 \Rightarrow x = A^T\lambda \Rightarrow Ax = AA^T\lambda \Rightarrow\Rightarrow (AA^T)^{-1}Ax = \lambda \Rightarrow \lambda = (AA^T)^{-1}b$，可以得到 $x = A^T(AA^T)^{-1}b$，对 $Ax = b$ ：

- m=n 时，方程具有唯一解
- m>n 时，方程无解，最小二乘解为 $\hat{x} = A^+b$，其中 A^+ 为 A 的伪逆矩阵
- m<n 时，方程有无穷解，最小二范数解为 $x = A^T(AA^T)^{-1}b$，其中 $A^T(AA^T)^{-1} = A^+$ 即为 A 的伪逆矩阵，证毕.

3 证明：$min\ ||X-YQ|| = min\ tr[(X-YQ)^T(X-YQ)] = min\ tr(X^TX - X^TYQ - Q^TY^TX + Q^TY^TYQ)$ 由 tr 的性质，$tr(X^TX - X^TYQ - Q^TY^TX + Q^TY^TYQ) = tr(X^TX) - 2tr(Q^TY^TX) + tr(Q^TY^TYQ) = tr(X^TX) - 2tr(Q^TY^TX) + tr(QQ^TY^TY) = tr(X^TX) - 2tr(Q^TY^TX) + tr(Y^TY)$，由于 $tr(X^TX), tr(Y^TY)$ 与 Q 无关，故原问题等价于：

$$max\ tr(Q^TY^TX)$$

令 $Z = V^TQ^TU$，z_{ii} 为 Z 的对角线元素，σ_{ii} 为 $\sum$ 的对角线元素，则

$$tr(\sum V^TQ^TU) = \sum_i \sigma_{ii}z_{ii} \le \sum_i \sigma_{ii}$$

当 Z=I 时等号成立，此时 $Q = UV^T$.

4 解：$Q(x) = x^TAx$，其中 $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$，$A = \begin{pmatrix} 7 & -4 & -2 \\ -4 & 1 & -4 \\ -2 & -4 & 7 \end{pmatrix}$ 作拉格朗日函数：

$$L(x) = x^TAx + \lambda(1 - x^Tx) \tag{1}$$

对其求偏导：

$$\frac{\partial L(x)}{\partial x} = (A^T + A)x - 2\lambda x = 0 \tag{2}$$

$$\frac{\partial L(x)}{\partial \lambda} = 1 - x^T x \tag{3}$$

可以得到：

$$(A - \lambda I)x = \begin{pmatrix} 7-\lambda & -4 & -2 \\ -4 & 1-\lambda & -4 \\ -2 & -4 & 7-\lambda \end{pmatrix} x = 0 \tag{4}$$

解得 $\lambda_1 = -3,\ \lambda_2 = \lambda_3 = 9$，带入 (4) 式可得当 $\lambda = -3$ 时，$x = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$，当 $\lambda = 9$ 时，$x = k_1 \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$. 又考虑到 $||x||_2^2 = 1$，当 $\lambda = -3$ 时，$Q(x) = -3$；当 $\lambda = 9$ 时，$Q(x) = 9$.

故当 $x = k_1 \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$，其中 k_1, k_2 满足 $||x||_2^2 = 1$ 时，$Q(x)$ 取得最大值 9.