۱) ماتریس $A=\begin{bmatrix}1 & 4\\ 1 & 0\end{bmatrix}$را به یک ماتریس متعامد یکه و یک ماتریس بالامثلثی تجزیه کنید.

۲) الف) در صورتیکه داشته باشیم، $Q^TQ = I$ ثابت کنید، Q حتما full column rank است. سپس نشان دهید $Q^+=Q^T$ (توجه کنید که Q لزوما معکوس پذیر نیست)

ب) در صورتیکه R یک ماتریس معکوس پذیر باشد و داشته باشیم، $AA^+=I$ (توجه کنید که A لزوما معکوس پذیر نیست). نشان دهید: $A^+=R^{-1}Q^+$

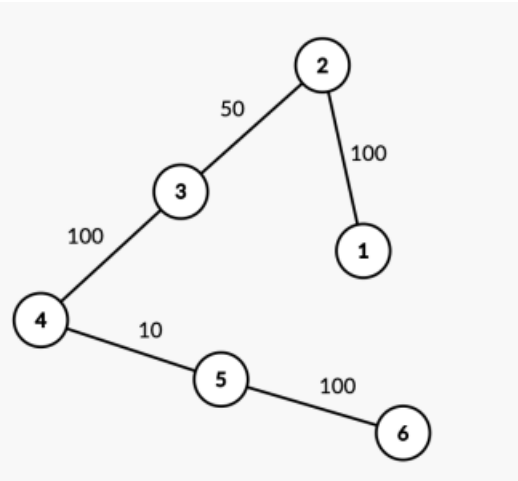
۳) فرض کنید، مقادیر متغیر b را به ازای مقادیر x_1 و x_2 اندازه گیری کرده ایم و مقادیر زیر را داریم.

x1	x2	b
1	2	1
0	1	2
1	-1	0

می‌خواهیم مقدار متغیر b را به صورت یک رابطه خطی از x_1 و x_2 ($b'=c_1x_1+c_2x_2$) تخمین بزنیم به طوریکه میزان خطای $\|b' - b\|_2$ مینیمم شود. مقادیر c_1، c_2 و را به نحوی که شرط بالا برقرار باشد، محاسبه کرده و میزان حداقل خطا را برای داده‌های بالا به دست آورید.

۴-

گراف وزندار زیر را در نظر بگیرید.



الف) ماتریس لاپلاسین را برای این گراف به دست آورید.

ب) در صورتیکه مقادیر، زیر بردار ویژه‌های یکه ماتریس لاپلاسین باشند و بخواهیم گراف را به دو کلاستر تقسیم کنیم. محاسبه کنید، کدام نودها در یک کلاستر قرار می‌گیرند (مراحل و محاسبات کامل نوشته شود.)

$[-0.56, -0.32, 0.43, 0.62, -0.06, -0.1]$ $[0.36, 0.34, 0.25, 0.19, -0.55, -0.6]$

$[0.36, -0.58, 0.61, -0.4, 0.04, -0.03]$ $[0.3, -0.33, -0.23, 0.3, -0.6, 0.55]$

$[-0.41, 0.41, 0.41, -0.41, -0.41, 0.41]$ $[0.41, 0.41, 0.41, 0.41, 0.41, 0.41]$

۵-

فرض کنید جدول زیر را به ازای اسناد d_1 تا d_6 داریم.

	d_1	d_2	d_3	d_4	d_5	d_6
ship	1	0	1	0	0	0
boat	0	1	0	0	0	0
ocean	1	1	0	0	0	0
drug	1	0	0	1	1	0
cancer	0	0	0	1	0	1

الف) در صورتیکه از tf برای وزندهی به کلمات هر سند استفاده کنیم. مشخص کنید، در پاسخ به پرسوجوی q_1 = boat drug رتبه‌بندی اسناد به ترتیب مرتبط‌ترین به چه شکلی خواهد بود؟ (روش vector space)

ب) فرض کنید، می دانیم اسناد را می‌توانیم در دو گروه قرار دهیم بنابراین ماتریس بالا را می‌خواهیم با یک ماتریس rank 2 تخمین بزنیم. و سپس با استفاده از آن ماتریس، رتبه بندی اسناد را به دست آوریم. این کار چه سودی خواهد داشت ؟ با توجه به داده های SVD که در ادامه داده شده اند، مرتبط‌ترین سند به q_1 کدام است و آنرا چگونه به دست می‌آوریم.

U:

1	2	3	4	5
−0.44	−0.30	0.57	0.58	0.25
−0.13	−0.33	−0.59	0.00	0.73
−0.48	−0.51	−0.37	0.00	−0.61
−0.70	0.35	0.15	−0.58	0.16
−0.26	0.65	−0.41	0.58	−0.09

Σ

2.16	0.00	0.00	0.00	0.00	0.00
0.00	1.59	0.00	0.00	0.00	0.00
0.00	0.00	1.28	0.00	0.00	0.00
0.00	0.00	0.00	1.00	0.00	0.00
0.00	0.00	0.00	0.00	0.39	0.00

V^T

	d_1	d_2	d_3	d_4	d_5	d_6
1	−0.75	−0.28	−0.20	−0.45	−0.33	−0.12
2	−0.29	−0.53	−0.19	0.63	0.22	0.41
3	0.28	−0.75	0.45	−0.20	0.12	−0.33
4	0.00	0.00	0.58	0.00	−0.58	0.58
5	−0.53	0.29	0.63	0.19	0.41	−0.22
6	0.00	0.00	0.00	−0.58	0.58	0.58

ج) اگر یک روش متن‌کاوی، برای پرس و جوی Q_1 تعداد ۱۰ سند مرتبط تشخیص داده باشد، و در واقعیت تعداد اسناد مرتبط به این پرسوجو ۵ سند باشد. مقدار precision و recall برای روش متن‌کاوی مذکور چقدر است؟