CSCI 5304 HW 4

Jingxiang Li

November 3, 2014

Problem 1

Write a 2×2 Givens rotation $\begin{bmatrix} c & s \\ -s & c \end{bmatrix}$ as a product of two Householder reflections. Hint: one way to do this involves making one of the two Householder reflectors represent the reflection across a coordinate axis.

Solution

Here we first consider a vector $v = (\cos(\theta), \sin(\theta))^T$ in $\mathbb{R}^2$

Let $c = \cos(\phi)$, $s = \sin(\phi)$, then $R = \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{bmatrix}$

$$Rv = \begin{bmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{bmatrix} \cdot \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} = \begin{bmatrix} \cos(\theta - \phi) \\ \sin(\theta - \phi) \end{bmatrix}$$

So that $R(\phi) : \theta \to \theta - \phi$

Then let's consider P as a 2×2 Householder rotation, where $P = \begin{bmatrix} \cos(\psi) & \sin(\psi) \\ \sin(\psi) & -\cos(\psi) \end{bmatrix}$

$$Pv = \begin{bmatrix} \cos(\psi) & \sin(\psi) \\ \sin(\psi) & -\cos(\psi) \end{bmatrix} \cdot \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} = \begin{bmatrix} \cos(\psi - \theta) \\ \sin(\psi - \theta) \end{bmatrix}$$

So that $P(\psi) : \theta \to \psi - \theta$

Then Let $Rv = P_2 P_1 v$, where P_2 is based on ψ_2 and P_1 is based on ψ_1, we have

$$\theta - \phi = \psi_2 - (\psi_1 - \theta)$$
$$\phi = \psi_1 - \psi_2$$

Then let $\psi_1 = \phi$, $\psi_2 = 0$, we have $Rv = P_2 P_1 v$, where

$$P_1 = \begin{bmatrix} c & s \\ s & -c \end{bmatrix} \quad P_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Problem 2

Prove or disprove: any square orthogonal matrix can be written as a product of Householder reflectors.

Proof

Consider A as a square orthogonal matrix in $\mathbb{R}^{n\times n}$. Since A is orthogonal then

$$||A||_2 = ||AI||_2 = |||I||_2 = 1$$

Then let P be a Householder reflection matrix, so that

$$A_1 = PA = \begin{bmatrix} 1 & w^T \\ 0 & B \end{bmatrix}$$

where w is a vector in $\mathbb{R}^{n-1}$, B is a square matrix in $\mathbb{R}^{(n-1)\times(n-1)}$

Now we prove that $w = 0$:

Notice that $||A_1||_2 = ||PA||_2 = 1$, $||A_1 \begin{bmatrix} 1 \\ w \end{bmatrix} ||_2 \geq 1 + w^T w$, so that

$$1 = ||A_1||_2 \geq 1 + w^T w$$

hence $w = 0$

Since $w = 0$, $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix}$, and since A_1 is still orthogonal, B must be an orthogonal square matrix in $\mathbb{R}^{(n-1)\times(n-1)}$, so that we can repeat this procedure until we transform A into an identity matrix. So now we have

$$\begin{aligned} P_{n-1} \dots P_2 P_1 A = I \\ A = (P_{n-1} \dots P_2 P_1)^{-1} \\ A = P_1 P_2 \dots P_{n-1} \end{aligned}$$

Q.E.D.

Problem 3

Prove or disprove: any square orthogonal matrix can be written as a product of Givens rotations.

Proof

Consider A as a square orthogonal matrix in $\mathbb{R}^{n\times n}$.

First, It's easily to find a series of Given rotation $R_{n-1}, \dots, R_2, R_1$ so that

$$A_1 = R_{n-1} \dots R_2 R_1 A = \begin{bmatrix} 1 & w^T \\ 0 & B \end{bmatrix}$$

And similar to the proof in problem 2, we know that $1 = ||A_1||_2 \geq ||A_1 \begin{bmatrix} 1 \\ w \end{bmatrix}||_2 \geq 1 + w^T w$, then $w = 0$,

$$A_1 = \begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix}$$

where B is a square orthogonal matrix in $\mathbb{R}^{(n-1)\times(n-1)}$, so that we can repeat this procedure until A is transformed into an Identity matrix.

$$\begin{aligned} R_k \dots R_2 R_1 A &= I \\ A &= (R_k \dots R_2 R_1)^{-1} \\ A &= R_1^T R_2^T \dots R_k^T \end{aligned}$$

Notice that k is determined by the number of rotations we apply, and if R is a Given rotation R^T is still a Given rotation. Q.E.D.

Problem 4

What are the eigenvalues and eigenvectors for a Householder reflector of the form $I - 2uu^T$ where $u^T u = 1$?

Solution

First, Let's consider u as an eigenvector, then $(I - 2uu^T)u = u - 2u = -u$ which indicates that the corresponding eigenvalue is -1. Then we consider vector v where $u^T v = 0$, then $(I - 2uu^T)v = v$, suggesting that v is an eigenvector with corresponding eigenvalue 1. Note that the dimension of subspace perpendicular to u is $n - 1$. Therefore except for u, all other $n - 1$ eigenvectors should be the basis of the subspace perpendicular to u, with corresponding eigenvalues 1.

Consider $(v_1, v_2, \ldots, v_{n-1})$ be the basis of the subspace perpendicular to u
the eigenvector matrix is

$$\begin{bmatrix} u & v_1 & v_2 & \cdots & v_{n-1} \end{bmatrix}$$

with the corresponding eigenvalues $(-1, 1, \ldots, 1)$

Problem 5

Let $P = \begin{bmatrix} c & s \\ s & -c \end{bmatrix}$ with $c^2 + s^2 = 1$. Prove or disprove: P is an orthogonal transformation. Is this a Givens rotation, a Householder transformation, or neither? If a Givens rotation, write it as a Givens rotation. If a Householder transformation, write it in the form $I - 2uu^T$ for some unit vector u.

Solution

First, P is orthogonal, since $P^T P = \begin{bmatrix} c^2 + s^2 & cs - cs \\ cs - cs & c^2 + s^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Then, it is not a Givens rotation, since Givens rotation matrix is not symmetric.

Lastly, it is a Householder transformation. Now we shall prove it:

Let $u = (\cos(\theta), \sin(\theta))^T$, then

$$I - 2uu^T = \begin{bmatrix} 1 - 2\cos^2(\theta) & -2\sin(\theta)\cos(\theta) \\ -2\sin(\theta)\cos(\theta) & 1 - 2\sin^2(\theta) \end{bmatrix} = \begin{bmatrix} \cos(\pi + 2\theta) & \sin(\pi + 2\theta) \\ \sin(\pi + 2\theta) & -\cos(\pi + 2\theta) \end{bmatrix}$$

Let $c = \cos(\pi + 2\theta)$, $s = \sin(\pi + 2\theta)$, $I - 2uu^T$ has the same form of P.
Let $\theta = (\pi - \arccos(c))/2$, $u = (\cos(\theta), \sin(\theta))^T$, then $P = I - 2uu^T$

Problem 6

Paul, Jane, and Ann, share information about their likes and dislikes of movies in order to make decisions about selecting films to see. They rates films they see with a scale of 0 to 10, (10 means they liked the movie very much). Here is the status of their table of ratings when Ann was interested in a new film which soon came to a 'theater near her' (titled 'Title 6' in the table):

Table 1: Problem 6

movie	Paul	Jane	Ann
Title-1	4	8	4
Title-2	9	3	8
Title-3	2	6	1
Title-4	7	4	4
Title-5	8	3	6
Title-6	3	8	x

Ann generally follows a combination of Paul and Jane's ratings. Ann wants to predict how well she will like the movie Title-6, which Paul and Jane have already seen. Ann reasons as follows: she will give her 'similarity' coefficients α and β for Paul and Jane respectively. If the missing rating (call it x) were known then the column of Ann's ratings should be the closest in the least-squares sense to the combination of Paul's and Jane's ratings:

Determine, α, β using the first 5 movie ratings, and from that the combination of Paul's and Jane's ratings that is closest to Ann's, in a least squares sense. Use the resulting best combination to infer the induced rating for Ann for Title-6. Should Ann see 'Title 6' ? Is her taste closer to Paul's or to Jane's?

Solution

Here we apply QR decomposition to solve the least square problem.

```
X = [4, 8; 9, 3; 2, 6; 7, 4; 8, 3];
y = [4; 8; 1; 4; 6];

[Q R] = qr(X);
Q = Q(:, 1:2);
R = R(1:2, :);
%% Here we use inv() since the dimension of R is very small
theta = inv(R) * (Q' * y);
alpha_0 = theta(1);
beta_0 = theta(2);
%% Make prediction
X_new = [3, 8];
prediction = X_new * theta;

alpha_0
% > alpha_0 =  0.77036

beta_0
% > beta_0 =  0.0093010

prediction
% > prediction =  2.3855
```

Here we see the predicted rating for Ann for Title-6 is 2.3855, suggesting that Ann should not see it. Since $\hat{\alpha} = 0.77036$, and $\hat{\beta} = 0.0093010$, her taste is closer to Paul's than Jane's.

Problem 7

The data in the file lsidata.m contains the term-frequency matrix A for a collection of 2340 documents, using a dictionary of 21839 words. Load this data into matlab and design a query to retrieve all documents containing the word "bipolar". Apply this query to the original term frequency matrix (with columns scaled to unit length) and then repeat this procedure using the best rank-50 approximation A 50 to the term-frequency matrix A. Only the document headlines are provided, but the word counts are based on the document contents which can be found at the web site mentioned within the datafile.

Hints: use svds to obtain the singular value decomposition $A = U\Sigma V^T$ instead of svd, because (a) the matrix A is too big for svd and (b) you can get the rank 50 approximation A 50 directly from svds. Do not try to form the rank-50 approximation explicitly - rather work directly with the factors U, Σ, V obtained from svds. Do not forget to normalize the columns to unit length before applying a query or computing the SVD. A skeleton matlab script for this problem is given in LSIpreamble.m.

The query vector is a vector with 1's in the positions corresponding to the words in your query and zeros for all other words. So if your query is only one word, then the query vector has only one nonzero element. If n is the number of words, and m the number of documents, then q is an n-vector with only a few nonzero elements. The term-frequency matrix A is an $n \times m$ with the j-th column corresponding to the j-th document. If each column a_j (for $j = 1, \ldots, m$) is normalized to length 1 in the usual 2-norm, and the query vector q is also normalized to unit length, then the inner product $q \cdot a_j$ is the cosine of the angle between the two vectors. If the vectors are almost the same, then the angle will be small and the cosine will be very close to 1. If all the columns of A are normalized in this fashion, then all the cosines can be computed at once with the formula $q^T A$. So your task is to compute $q^T A$ and also $q^T A_{50}$. Once you have computed these two vectors of similarities, you need to sort each of them to find the positions of the 10 biggest values, and then print out the document headlines corresponding to those positions.

You will need to use the sort function in matlab, which returns two results. The first result is the values sorted, and the second result consists of the indices of the positions of those values. You can use that second result to retrieve the document headlines. A sample script which carries out all of this on a toy example is given in LSItoy.m.

Solution

```
clear; close all; clc
% Read Data
lsidata;
[n, m] = size(A);
k = 50;

% Now Calculate A_50 Using svd
[U, S, V] = svds(A, k);
A_50 = U(:, 1:k) * S(1:k, 1:k) * V(:, 1:k)';

% q is the query vector in the sparse form
q_tmp = [1897, 1, 1; n, 1, 0];
q = spconvert(q_tmp);

% result_origin
result_origin = q' * A;
% Only 112th row has non-zero output, suggesting that the query result
    is the 112th document
headlines(112)
% > 112 Smoking Alters Brain Chemical

% result_reduced
result_reduced = q' * A_50;
[s index] = sort(result_reduced, 'descend');
headlines(index(1:10))

% > 492 Ultrasound Reveals Thinking Brain
% > 240 Enzyme Linked To Brain Aneurysm
% > 176 Scans Reveal Brain Reacting to Cocaine
% > 112 Smoking Alters Brain Chemical
% > 423 Diet Drugs Affect Brain Cells
% > 177 Leptin: Possible Diabetes Treatment?
% > 448 Diet Drugs Affect Brain Cells
% > 131 Obesity Hormone Regulates Blood Sugar
% > 392 Brain Chemicals Mimic Marijuana
% > 251 Newborn Brain Link to Mental Ills
```

Table 2 contains the query result based on the original document-word frequency matrix, Table 3 is the result based on the reduced matrix.

Table 2: Query result based on the original matrix

Index	Headline
112	Smoking Alters Brain Chemical

Table 3: Query result based on the reduced matrix

Index	Headline
492	Ultrasound Reveals Thinking Brain
240	Enzyme Linked To Brain Aneurysm
176	Scans Reveal Brain Reacting to Cocaine
112	Smoking Alters Brain Chemical
423	Diet Drugs Affect Brain Cells
177	Leptin: Possible Diabetes Treatment?
448	Diet Drugs Affect Brain Cells
131	Obesity Hormone Regulates Blood Sugar
392	Brain Chemicals Mimic Marijuana
251	Newborn Brain Link to Mental Ills