CS4220 PS 2

Due: Mon, February 2, 2015

1.(Ascher and Greif, Section 4.6, Problem 5) Determine whether the following statement is true or false: If A is non-singular, then for any induced norm, $\|A^{-1}\| = \|A\|^{-1}$.

Solution: The statement is false. For a non-singular matrix A, we have $AA^{-1} = I$. The consistency condition of the matrix multiplication requires $\|AB\| \leq \|A\|\|B\|$. We can have

$$\|I^2\| = \|I\| \leq \|I\| \cdot \|I\|, \quad \text{so} \quad \|I\| \geq 1.$$

and

$$\|I\| = \|AA^{-1}\| \leq \|A\| \cdot \|A^{-1}\|,$$

so

$$\|A^{-1}\| \geq \frac{\|I\|}{\|A\|} \geq \frac{1}{\|A\|} = \|A\|^{-1}.$$

We can conclude the equality statement is false in general.

A quick MATLAB experiment below demonstrates the statement is false for all the following matrix norms and its inverse matrix norms of an random matrix A: $\|A\|_1$, $\|A\|_2$ and $\|A\|_\infty$.

```
>> A=rand(4,4)
A =
    0.7482    0.9133    0.9961    0.9619
    0.4505    0.1524    0.0782    0.0046
    0.0838    0.8258    0.4427    0.7749
    0.2290    0.5383    0.1067    0.8173

>>norm(A,1)
ans = 2.5587
>>norm(inv(A), 1)
ans = 7.9626

>> norm(A,2)
ans = 2.3476
>> norm(inv(A), 2)
ans = 5.8718

>> norm(A,inf)
ans = 3.6195
>> norm(inv(A), inf)
ans = 8.7548
```

2.(Ascher and Greif, Section 4.6, Problem 6) For an $m \times n$ matrix A show that

$$\|A\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

Proof. Let matrix A be written as a collection of n column vectors:

$$A = \begin{pmatrix} \mathbf{a_1} & \mathbf{a_n} & \dots & \mathbf{a_n} \end{pmatrix},$$

and $\mathbf{x}$ be any vector with $\|\mathbf{x}\|_1 = 1$, thus we have $\sum_{j=1}^{n} |x_j| \le 1$. Using the definition of the vector ℓ_1-norm, we can proceed in the following steps to find an upper bound of $\|A\mathbf{x}\|_1$:

$$\begin{aligned}
\|A\mathbf{x}\|_1 &= \left\| \sum_{j=1}^{n} x_j \, \mathbf{a_j} \right\|_1 \le \sum_{j=1}^{n} \|x_j \mathbf{a_j}\|_1 \quad \text{(vector norm's triangle inequality)} \\
&= \sum_{j=1}^{n} |x_j| \; \|\mathbf{a_j}\|_1 \quad \text{(vector norm's scaling property: } \|x_j \mathbf{a_j}\|_1 = |x_j| \|\mathbf{a_j}\|_1) \\
&\le \sum_{j=1}^{n} |x_j| \left(\max_{1 \le j \le n} \|\mathbf{a_j}\| \right) \quad \text{(obviously } \|\mathbf{a_j}\|_1 \le \max_{1 \le j \le n} \|\mathbf{a_j}\|_1 \text{ for all } j) \\
&\le \max_{1 \le j \le n} \|\mathbf{a_j}\|_1 \quad \text{(we know } \sum_{j=1}^{n} |x_j| \le 1).
\end{aligned}$$

So we now have

$$\|A\|_1 = \max_{\|\mathbf{x}\|_1 = 1} \|A\mathbf{x}\|_1 \le \max_{1 \le j \le n} \|\mathbf{a_j}\|_1. \tag{1}$$

The equality of the equation above can be achieved by choosing $\mathbf{x} = \mathbf{e_j}$, where column j of A is the column producing the maximum value of $\|\mathbf{a_j}\|_1$, for all j:

$$\|A\mathbf{x}\|_1 = \|A\mathbf{e_j}\|_1 = \max_{1 \le j \le n} \|\mathbf{a_j}\|_1 \ge \max_{1 \le j \le n} \|\mathbf{a_j}\|_1. \tag{2}$$

Combining Eq.(1) and Eq.(2), we get

$$\max_{1 \le j \le n} \|\mathbf{a_j}\|_1 \le \|A\mathbf{x}\|_1 \le \|A\|_1 \le \max_{1 \le j \le n} \|\mathbf{a_j}\|_1. \tag{3}$$

Eq.(3) implies equality of the equation. By substituting the definition of vector ℓ_1-norm

$$\|\mathbf{a_j}\|_1 = \sum_{i=1}^{m} |a_{ij}|$$

into Eq.(3), we have proved that $\|A\|_1$ is equal to *the maximum absolute column sum* of the matrix:

$$\|A\|_1 = \max_{1 \le j \le n} \|\mathbf{a_j}\|_1 = \max_{1 \le j \le n} \sum_{i=1}^{m} |a_{ij}|.$$

$\square$

4.(Ascher and Greif, Section 4.6, Problem 7) Let A be symmetric positive definite (SPD). Show that the so-called energy norm

$$\|\mathbf{x}\|_A = \sqrt{\mathbf{x}^T A\mathbf{x}}$$

is indeed a vector norm.

Proof. A is a symmetric matrix if and only if A is orthogonally diagonalizable. That is, by using the principal axis theorem (see below), we know there exists an orthogonal matrix P such that $P^T AP = D$, where D is a diagonal matrix. Furthermore, the diagonal entries of D are eigenvalues of A that correspond, respectively. to the eigenvectors in P. Finally, all eigenvalues of a symmetric positive definite matrix are positive.

For any $\mathbf{x} \in \mathbb{R}^n$, we introduce a change of variable transformation

$$\mathbf{x} = P\mathbf{y}, \quad \mathbf{y} = P^{-1}\mathbf{x}, \tag{4}$$

where P is the orthogonal matrix which makes $P^T AP = D$. In fact, $P^{-1} = P^T$. Substituting $\mathbf{x} = P\mathbf{y}$ in the quadratic form $\mathbf{x}^T A\mathbf{x}$, we have

$$\mathbf{x}^T A\mathbf{x} = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T(P^T AP)\mathbf{y} = \mathbf{y}^T D\mathbf{y} = \sum_{j=1}^{n} \lambda_j y_j^2 > 0. \tag{5}$$

Because $d_{jj} = \lambda_j > 0$ in the diagonal matrix D, $D^{\frac{1}{2}}$ is always available:

$$D = D^{\frac{1}{2}} D^{\frac{1}{2}}. \tag{6}$$

With Eq.(5) and Eq.(6), the energy norm of any vector $\mathbf{x} \in \mathbb{R}^n$ for a symmetric positive definite matrix A can be rewritten as

$$\|\mathbf{x}\|_A = \sqrt{\mathbf{x}^T A\mathbf{x}} = \sqrt{\mathbf{y}^T D\mathbf{y}} = \sqrt{\mathbf{y}^T D^{\frac{1}{2}} D^{\frac{1}{2}}\mathbf{y}} = \|D^{\frac{1}{2}}\mathbf{y}\|_2.$$

So $\|x\|_A$ corresponds to $\|D^{\frac{1}{2}}\mathbf{y}\|_2 = \|D^{\frac{1}{2}}(P^{-1}x)\|_2$, which is an ℓ_2-norm. Therefore the energy norm associated with an SPD matrix A is indeed a vector norm.

Theorem (Principle Axis Theorem)**.** *Let A be an $n \times n$ symmetric matrix, then there is an orthogonal change of variable, $\mathbf{x} = P\mathbf{y}$, that transforms the quadratic form $\mathbf{x}^T A\mathbf{x}$ into a quadratic form $\mathbf{y}^T D\mathbf{y}$ with no cross-product term.*

□

4.(Ascher and Greif, Section 4.6, Problem 9) Find all values of a and b such that

$$A = \begin{pmatrix} a & 1 & 1+b \\ 1 & a & 1 \\ 1-b^2 & 1 & a \end{pmatrix}$$

is symmetric and positive definite.

Solution: Assume $a, b \in \mathbb{R}$. If A is symmetric, then $A = A^T$, so $1 - b^2 = 1 - b$, we must have $b = 1$ or 0. If A is positive definite, all the eigenvalues of A must be positive:

- When $b = 0$, solve the following characteristic equation for eigenvalues:

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} a-\lambda & 1 & 1 \\ 1 & a-\lambda & 1 \\ 1 & 1 & a-\lambda \end{vmatrix} = -(\lambda - a)^3 + 2 + 3(\lambda - a) \\
&= [\lambda - (a-1)]^2 [\lambda - (a+2)] = 0.
\end{aligned}$$

  $\lambda_1 = a + 2$, $\lambda_2 = \lambda_3 = a - 1$

- When $b = 1$, solve the following characteristic equation for eigenvalues:

$$\begin{aligned}
\det(A - \lambda I) &= \begin{vmatrix} a-\lambda & 1 & 0 \\ 1 & a-\lambda & 1 \\ 0 & 1 & a-\lambda \end{vmatrix} = -(\lambda - a)^3 + 2(\lambda - a) \\
&= (\lambda - a)[\lambda - (a + \sqrt{2})][\lambda - (a - \sqrt{2})] = 0.
\end{aligned}$$

  $\lambda_1 = a + \sqrt{2}$, $\lambda_2 = a$, $\lambda_3 = a - \sqrt{2}$.

Therefore, in order that A is symmetric and positive definite (all $\lambda > 0$), possible values of a and b are:

1. $b = 0$ and $a > 1$, or
2. $b = 1$ and $a > \sqrt{2}$.

4.(Ascher and Greif, Section 4.6, Problem 10) (a) Show that the matrix

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix}$$

is orthogonal $c^2 + s^2 = 1$.

(b) Givens rotations are based on rotation operations of the form[1]

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix}^T \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \tag{7}$$

- Use orthogonality to express α in terms of a_1 and a_2.
- Find c and s that do the job.

[1] The textbook of Ascher and Greif does not have the transpose on the system matrix in Eq.(7).

Solution:

(a) Let us call the Givens matrix $G = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$. We have

$$G^T G = \begin{pmatrix} c^2 + s^2 & 0 \\ 0 & c^2 + s^2 \end{pmatrix}.$$

Obviously, if $c^2 + s^2 = 1$, then the matrix G is orthogonal: $G^T G = I$.

(b) We pre-multiply the rotation equation by G to obtain

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = G \begin{pmatrix} \alpha \\ 0 \end{pmatrix} \implies \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} c\alpha \\ -s\alpha \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Therefore we have

$$\alpha = \frac{a_1}{c} = -\frac{a_2}{s}.$$

With any given vector $(a_1, a_2)^T$, we solve the following equations for c and s:

$$-\frac{a_1}{a_2} = \frac{c}{s}, \quad \text{if } a_2 \neq 0, \tag{8}$$

$$1 = c^2 + s^2. \tag{9}$$

So

$$s = \frac{1}{\sqrt{1 + (a_1/a_2)^2}}, \quad c = -\frac{(a_1/a_2)}{\sqrt{1 + (a_1/a_2)^2}}.$$

The c and s that do the job of the Givens rotations can be written in more compact form:

$$\text{if } a_2 \neq 0, \quad \text{let } r = -\frac{a_1}{a_2}, \quad c = \frac{r}{\sqrt{1+r^2}} \quad \text{and} \quad s = \frac{1}{\sqrt{1+r^2}};$$

and

$$\text{if } a_2 = 0, \quad \text{then } c = 1, \quad s = 0.$$

Note: It is straightforward to use the above results to show

$$\alpha^2 = \frac{a_1^2}{c^2} = a_2^2(1 + r^2) = a_1^2 + a_2^2.$$

Therefore the transformation keep the ℓ_2-norm of the vectors the same.

2: Recognizing rank Consider the MATLAB fragment

```
function [y] = ps2mult(x)
  n = length(x);
  A = reshape(1:n^2, n, n);
  y = A*x;
```

1. What is A for $n = 3$?
2. Show that A has rank two (independent of n).
3. Rewrite `ps2mult` so that it runs in $O(n)$ time.

Solution: For $n = 3$,

$$A = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

For the general $n \times n$ case of A the column vectors of A, called them $\mathbf{a_1}, \mathbf{a_2}, \ldots, \mathbf{a_n}$, are shown below:

$$A = (\mathbf{a_1} \quad \mathbf{a_2} \quad \ldots \quad \mathbf{a_n}) = \begin{pmatrix} 1 & n+1 & 2n+1 & . & . & . & (n-1)n+1 \\ 2 & n+2 & 2n+2 & . & . & . & (n-1)n+2 \\ 3 & n+3 & 2n+3 & . & . & . & (n-1)n+3 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ n & n+n & 2n+n & . & . & . & (n-1)n+n \end{pmatrix} \tag{10}$$

We can see obviously in Eq.(10) all column vectors of A are actually linear combinations of only two linearly independent vectors, which are

$$\mathbf{a_1} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ . \\ . \\ n \end{pmatrix} \quad \text{and} \quad \mathbf{w} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ . \\ . \\ 1 \end{pmatrix}. \tag{11}$$

Consequently, the column space of matrix A is spanned by these two vectors:

$$\mathrm{Col}\,(A) = \mathrm{span}\{\mathbf{a_1}, \mathbf{w}\}.$$

Rank of A is the number of linearly independent vectors in Col(A). Here the rank of this matrix A is 2, which is independent of n.

Matrix vector multiplication $A\mathbf{x}$ can be interpreted as the linear combination of column vectors of A with weights from components of vector $\mathbf{x} = (x_1\ x_2\ \ldots\ x_n)^T$:

$$A\mathbf{x} = x_1\mathbf{a_1} + x_2\mathbf{a_2} + \ldots + x_n\mathbf{a_n} = \sum_{j=1}^{n} x_j\mathbf{a_j}\,. \tag{12}$$

In order to reduce the number of flops in the multiplication operation of this special matrix A ($n \times n$) with any given vector $\mathbf{x}$ ($n \times 1$), we can exploit the structure of A as shown in Eq.(10) by expressing the column vectors of A as:

$$
\begin{aligned}
\mathbf{a_1} &= \mathbf{a_1} \\
\mathbf{a_2} &= \mathbf{a_1} + n\mathbf{w} \\
\mathbf{a_3} &= \mathbf{a_1} + 2n\mathbf{w} \\
\mathbf{a_4} &= \mathbf{a_1} + 3n\mathbf{w} \\
\ldots & \quad \ldots \\
\mathbf{a_n} &= \mathbf{a_1} + (n-1)n\mathbf{w}
\end{aligned}
$$

Substituting the these column vectors into Eq.(12), we get

$$
\begin{aligned}
A\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{a_j} &= x_1(\mathbf{a_1}) + x_2(\mathbf{a_1} + n\mathbf{w}) + \ldots + x_n(\mathbf{a_1} + (n-1)n\mathbf{w}) & (13) \\
&= \left(\sum_{j=1}^{n} x_j \right) \mathbf{a_1} + \left(\sum_{j=1}^{n} (j-1)x_j \right) n\,\mathbf{w} & (14)
\end{aligned}
$$

Now the number of operations in Eq.(14) is a total of $O(n)$ flops, instead of $O(n^2)$ as required by general matrix vector multiplications.

Notes about code vectorization for Eq.(14) in MATLAB:

- on the right hand side of Eq.(14), $\sum_{j=1}^{n} x_j$ can be vectorized in MATLAB by `sum(x)`;
- $\sum_{j=1}^{n}(j-1)x_j$ is equal to the following dot product: $(\mathbf{a_1} - \mathbf{w})^T \cdot \mathbf{x}$, which can be vectorized in MATLAB as `(a1 -1)'*x`;
- vectorization in MATLAB eliminates the need to create vector $\mathbf{w}$ for implementing Eq.(14). (See the last line in function `ps2mult`.)

```
function [y] = ps2mult(x)
%
% x is an (n by 1) vector
% A = reshape(1:n^2, n, n)
%
% This routine is to return y=A*x in O(n) flops
%
  n=length(x);
%
% the matrix  A = reshape(1:n^2, n, n);
%
  a1=[1:n]';
%  w = ones(n,1);
%
% Recognizing A (n by n) is a rank 2 matrix whose column
% space is span{ a_1, w } (a1 and w are defined as above):
%
% A = [a1 a1+n*w a1+2n*w .... a1+(n-1)n*w ]
%
% We exploit the structure of matrix A here so that  A*x
% only uses O(n) flops, instead of O(n^2) flops.
%
%
  sum1=0;
  sum2=0;
%
%  for j=1:n
%    sum1=sum1+x(j);
%    sum2=sum2+(j-1)*x(j);
%  end
% Use vectorized operations in Matlab instead of the loop:
  sum1 = sum(x);
  sum2 = (a1-1)'*x;
%
% remember to multiply sum2 by n below (according to the simple formula
% we derived in the solution):
%
  y = sum1*a1 + n*sum2;
```