

HW5

Thursday, March 7, 2024 1:38 PM

5.3. Consider the matrix

$$A = \begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix}.$$

- (a) Determine, on paper, a real SVD of A in the form $A = U\Sigma V^T$. The SVD is not unique, so find the one that has the minimal number of minus signs in U and V .
- (b) List the singular values, left singular vectors, and right singular vectors of A . Draw a careful, labeled picture of the unit ball in \mathbb{R}^2 and its image under A , together with the singular vectors, with the coordinates of their vertices marked.
- (c) What are the 1-, 2-, ∞ -, and Frobenius norms of A ?
- (d) Find A^{-1} not directly, but via the SVD.
- (e) Find the eigenvalues λ_1, λ_2 of A .
- (f) Verify that $\det A = \lambda_1 \lambda_2$ and $|\det A| = \sigma_1 \sigma_2$.
- (g) What is the area of the ellipsoid onto which A maps the unit ball of \mathbb{R}^2 ?

$$(a) AA^T = U\Sigma V^T (U\Sigma V^T)^T = U\Sigma \Sigma U^T$$

$$AA^T = \begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix} \begin{bmatrix} -2 & -10 \\ 11 & 5 \end{bmatrix} = \begin{bmatrix} 125 & 75 \\ 75 & 125 \end{bmatrix}$$

$$\begin{vmatrix} 125-\lambda & 75 \\ 75 & 125-\lambda \end{vmatrix} = (125-\lambda)^2 - 75^2 = \lambda^2 - 250\lambda + 10000$$

$$\Rightarrow \lambda_2 = 50 \text{ and } \lambda_1 = 200$$

$$\Rightarrow \sigma_2 = 5\sqrt{2} \text{ and } \sigma_1 = 10\sqrt{2}$$

$U^T AA^T U = \Sigma^2$. We can see U are the eigenvectors of AA^T .

$$N\left(\begin{bmatrix} 75 & 75 \\ 75 & 75 \end{bmatrix}\right) = \begin{bmatrix} u_2 \\ -1 \\ 1 \end{bmatrix} \quad N\left(\begin{bmatrix} -75 & 75 \\ 75 & -75 \end{bmatrix}\right) = \begin{bmatrix} u_1 \\ 1 \\ 1 \end{bmatrix}$$

$$\text{Thus } U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

$$A^T A = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^2 V^T.$$

Thus V is the eigenvectors of $A^T A$.

$$\begin{bmatrix} -2 & -10 \\ 11 & 5 \end{bmatrix} \begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix} = \begin{bmatrix} 104 & -72 \\ -72 & 146 \end{bmatrix}$$

$$\begin{vmatrix} 104-\lambda & -72 \\ -72 & 146-\lambda \end{vmatrix} = (104-\lambda)(146-\lambda) - 72^2 = \lambda^2 - 250\lambda + 10000$$

which is the same as $A A^T$

$$N\left(\begin{bmatrix} -96 & -72 \\ -72 & -54 \end{bmatrix}\right) = N\left(\begin{bmatrix} -16 & -72 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$N\left(\begin{bmatrix} 54 & -72 \\ -72 & 96 \end{bmatrix}\right) = N\left(\begin{bmatrix} 54 & -72 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

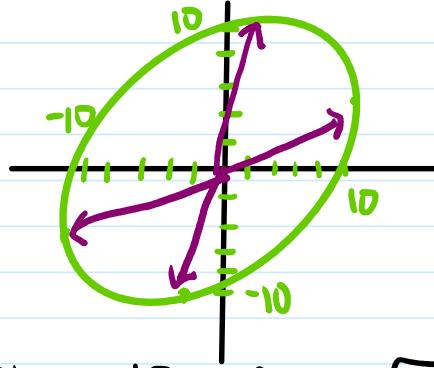
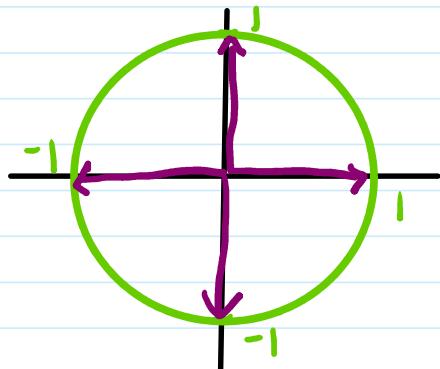
$$V = \frac{1}{5} \begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix}$$

$$\text{Thus } A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 10\sqrt{2} & 0 \\ 0 & 5\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & -3/4 \\ 1 & 3/4 \end{bmatrix} \begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix}$$

$$b) \sigma_1 = 10\sqrt{2} \quad v_1 = \pm \frac{1}{5} \begin{bmatrix} -3 \\ 4 \end{bmatrix} \quad v_2 = \pm \frac{1}{5} \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

$$u_1 = \pm \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad u_2 = \pm \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$Ae_1 = \begin{bmatrix} -2 \\ -10 \end{bmatrix} \quad Ae_2 = \begin{bmatrix} 11 \\ 5 \end{bmatrix}$$



$$c) \|A\|_1 = 16 \quad \|A\|_2 = 10\sqrt{2} \quad \|A\|_\infty = 15 \quad \|A\|_F = \sqrt{200+50} = 5\sqrt{10}$$

$$d) A^{-1} = (U \Sigma V^T)^{-1} = V \Sigma^{-1} U^T$$

$$= \frac{1}{5} \begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{10\sqrt{2}} & 0 \\ 0 & \frac{1}{5\sqrt{2}} \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$= \frac{1}{50} \begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{50} \begin{bmatrix} -3 & 4 \\ 4 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

$$= \frac{1}{50} \begin{bmatrix} 5/2 & 1/2 \\ 5 & -1 \end{bmatrix}$$

$$e) | -2 - \lambda \quad 11 | = (-2 - \lambda)(5 - \lambda) + 110 = \lambda^2 - 3\lambda + 100 = 0$$

e) $\begin{vmatrix} -2-\lambda & 1 \\ -10 & 5-\lambda \end{vmatrix} = (-2-\lambda)(5-\lambda) + 110 = \lambda^2 - 3\lambda + 100 = 0$
 $\lambda_{1,2} = \frac{3 \pm \sqrt{9-400}}{2} = \frac{3 \pm i\sqrt{391}}{2}$

f) $\det A = -10 + 110 = 100$

$$\lambda_1, \lambda_2 = \left(\frac{3+i\sqrt{391}}{2} \right) \left(\frac{3-i\sqrt{391}}{2} \right) = \frac{9+391}{4} = 100$$

$$\sigma_1, \sigma_2 = (10\sqrt{2})(5\sqrt{2}) = 50 \cdot 2 = 100 = 100\pi$$

g) Area = $\pi |Ae|$; $|Ae| = \pi \det A = 100\pi$

2. [Q2] (10 points) (a) If $A \in \mathbb{R}^{m \times n}$ and $E \in \mathbb{R}^{m \times n}$, show that

$$\sigma_{\max}(A+E) \leq \sigma_{\max}(A) + \|E\|_2 \quad \text{and} \quad \sigma_{\min}(A+E) \geq \sigma_{\min}(A) - \|E\|_2.$$

Comment on the (absolute) condition number of $\|A\|_2$ as a function of A .

(b) If $A \in \mathbb{R}^{m \times n}$, $m > n$ and $z \in \mathbb{R}^m$, show that

$$\sigma_{\max}([A \ z]) \geq \sigma_{\max}(A) \quad \text{and} \quad \sigma_{\min}([A \ z]) \leq \sigma_{\min}(A).$$

2a) Note $\|A\|_2^2 = p(A) \Rightarrow \|A\|_2 = \sigma_{\max}(A)$

By the triangle inequality $\sigma_{\max}(A+E) = \|A+E\|_2 \leq \|A\|_2 + \|E\|_2$. $\|A\|_2$ is proportional to $\|E\|_2$

$$\begin{aligned} \|A+E\|_2 &= \|A + E\|_2 \leq \|A\|_2 + \|E\|_2 \\ &= \sigma_{\max}(A) + \|E\|_2 \end{aligned}$$

By the reverse triangle inequality

$$|\sigma_{\max}(A) - \|E\|_2| = |\|A\|_2 - \|E\|_2| \leq \|A+E\|_2$$

$$= \sigma_{\max}(A+E)$$

b) Since if $z \notin \text{col}(A)$, then

$$\text{rank}([A \ z]) > \text{rank}(A) \text{ so } \exists! \lambda' \in \lambda([A \ z]) \ni \lambda' \notin \lambda(A)$$

Thus $\sigma_{\max}([A \ z]) \geq \sigma_{\max}(A)$. Since $\sqrt{\lambda}$ could be greater than $\sigma_{\max}(A)$

Similarly, $\sqrt{\lambda}$ could be less than $\sigma_{\min}(A)$

$$\Rightarrow \sigma_{\min}([A \ z]) \leq \sigma_{\min}(A)$$

3. [Q3] (10 points) (a) Show that if $A \in \mathbb{R}^{m \times n}$, then $\|A\|_F \leq \sqrt{\text{rank}(A)} \|A\|_2$.

(b) Show that if $A \in \mathbb{R}^{m \times n}$ has rank n , then $\|A(A^T A)^{-1} A^T\|_2 = 1$.

a) Let $\text{rank} A = k$ Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$

$$\begin{aligned} k\|A\|_2^2 &= k(\lambda_1)^2 \geq \sum_{i=1}^k (\lambda_i)^2 = \sum_{i=1}^k \sigma_i^2 \\ &= \|A\|_F^2 \end{aligned}$$

Thus $\|A\|_F \leq \sqrt{k} \|A\|_2$

$$b) (A(A^T A)^{-1} A^T)^2 = A(A^T A)^{-1} \cancel{A^T} \cancel{A} (A^T A)^{-1} A^T$$

Thus $A(A^T A)^{-1} A^T$ is a projection matrix $\therefore \|A(A^T A)^{-1} A^T\|_2 = 1$

4. [Q4] (a) (10 points) Given $A \in \mathbb{R}^{n \times n}$, let $A = U\Sigma V^T$ be an SVD of A , where $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$. Let $B = [U\text{diag}(1, \dots, 1, -1)V^T]$ such that $\det(B) = -\det(A)$ and $\|A - B\|_F = 2\sigma_n$. Show that for any singular values $\sigma_1, \sigma_2, \dots, \sigma_{n-1} (\geq \sigma_n)$, there exists $C \in \mathbb{R}^{n \times n}$ such that $\det(C) = \det(B) = -\det(A)$, and $\|A - C\|_F < \|A - B\|_F = 2\sigma_n$.

(Hint: to construct C , modify σ_n and σ_{n-1} of A only. Change the sign of one and keep the other, change the values of the two, and make sure that the product of the two modified scalars does not change in absolute value.)

6.1. If P is an orthogonal projector, then $I - 2P$ is unitary. Prove this algebraically, and give a geometric interpretation.

Let U_C and V_C be the same U and V

from A . Let $\Sigma_C = \begin{bmatrix} \sigma_1 & & \\ \vdots & \ddots & 0 \\ 0 & \cdots & \sigma_{n-1} \sigma_n & -1 \end{bmatrix}$

Note: $\det C = \det(U \Sigma_C V^T) = \det U \det \Sigma_C \det V = \det A$.

$$\Sigma \cdot \Sigma_C = \begin{bmatrix} \sigma_1 & & & 0 \\ 0 & \sigma_2 & & \\ 0 & 0 & \sigma_{n-1} & -\sigma_{n-1} \sigma_n \\ 0 & 0 & 0 & \sigma_{n+1} \end{bmatrix} \quad \|A - C\|_F = \|\Sigma - \Sigma_C\|_F$$

$$\frac{\|\Sigma - \Sigma_C\|_F^2}{\sigma_n^2} = \frac{(\sigma_n + 1)^2 + \sigma_{n-1}^2(1 - \sigma_n)^2}{\sigma_n^2} = 1 + \frac{2}{\sigma_n} + \frac{1}{\sigma_n^2} + \sigma_{n-1}^2 \left(\frac{1}{\sigma_n^2} - \frac{2}{\sigma_n} + 1 \right)$$

Let $\sigma_n \geq 1$

$$1 + \frac{2}{\sigma_n} + \frac{1}{\sigma_n^2} + \sigma_{n-1}^2 \left(\frac{1}{\sigma_n^2} - \frac{2}{\sigma_n} + 1 \right) \leq 1 + 2 + 1 + \sigma_{n-1}^2(1 - 2 + 1) = 3 \leq 4.$$

It follows,

$$\|A - C\|_F^2 \leq 4\sigma_n^2 \Rightarrow \|A - C\|_F \leq 2\sigma_n.$$

Let $0 < \sigma_n \leq 1$

$$\begin{aligned} & (\sigma_n^2 + 2\sigma_n + 1 + \sigma_{n-1}^2(1 - 2\sigma_n + \sigma_n^2))\sigma_n^2 \\ & \leq (1 + 2 + 1 + \sigma_{n-1}^2(1 - 2 + 1))\sigma_n^2 \end{aligned}$$

$$\leq (1+2+1+\sigma_{n-1}^2(1-2+1))\sigma_n^2 \\ = 3\sigma_n^2 \leq 4\sigma_n^2$$

b) $(I-2P)(I-2P) = I - 2P - 2P + 4P^2 = I - 4P + 4P = I$

Geometrically, similarly to Householders, $I-2P$ reflects across the subspace spanned by the projected vectors.

5. [Q5] (10 points) (a) Implement the Golub-Kahan (GK) bidiagonalization of a matrix. Test it on $F \in \mathbb{R}^{10 \times 10}$ obtained as follows

```
rgn('default');
F = randn(10,10);
```

Make sure that your bidiagonal matrix has the same singular values as F .

- (b) Generate a matrix $A \in \mathbb{R}^{(1024^2+1) \times 32}$ as follows

```
col = linspace(-1,1,1024*1024+1)';
A = col.^{0:31};
```

Apply Householder QR to A and get $R \in \mathbb{R}^{32 \times 32}$, then apply GK to R and get bidiagonal $B \in \mathbb{R}^{32 \times 32}$ (no need to retrieve Q for this problem). Compute the 5 largest and 5 smallest singular values of A from the eigenvalues of $\begin{bmatrix} 0 & B^T \\ B & 0 \end{bmatrix}$. Compare these singular values with those computed by taking the square root of the 5 largest and 5 smallest eigenvalues of $A^T A$. What conclusion do you draw? Is it a good idea

to compute the eigenvalues of $\begin{bmatrix} 0 & A^T \\ A & 0 \end{bmatrix}$ directly, and why?

a) Same σ ;

b) $\begin{bmatrix} 0 & B^T \\ B & 0 \end{bmatrix}$ has closer values $A^T A$ did not calculate smaller σ ; very well

$\begin{bmatrix} 0 & A^T \\ A & 0 \end{bmatrix}$ Is a bad idea because it is way too large.

```

%Problem 5a
rng('default');
F = randn(10,10);

K=HBD(F);
Sig=iterGKbdiag(K);
[U,S,V] = svd(K);

%Problem 5b
col = linspace(-1,1,1024*1024+1)';
A = col.^ (0:31);

R=HQR(A);
B=HBD(R);

[m,n]=size(B);

block = [zeros(m,n) B';B zeros(n,m)];
eblock = eig(block);
eblock = sort(eblock);
eblock = eblock(n+1:2*n);
sblock = sqrt([eblock(1:5);eblock(n-4:n)]);

e = eig(A'*A);
e = sort(e);
s = sqrt([e(1:5);e(n-4:n)]);

function R = HQR(A)
[m,n] = size(A);
R = A;

for k = 1:n
    x = R(k:m,k);
    e = zeros(length(x),1);
    e(1) = norm(x);
    if x(1) == 0
        beta = 1;
    else
        beta = sign(x(1));
    end
    u = beta*e + x;
    v = u / norm(u);
    R(k:m,k:n) = R(k:m,k:n) - 2*v*(v'*R(k:m,k:n));
end

R = R(1:n,1:n);
end

function R = HBD(A)
[m,n] = size(A);
R = A;

```

```

for k = 1:n
    %Left Hand Side
    x = R(k:m,k);
    e = zeros(length(x),1);
    e(1) = norm(x);
    if x(1) == 0
        beta = 1;
    else
        beta = sign(x(1));
    end
    u = beta*e + x;
    v = u / norm(u);
    R(k:m,k:n) = R(k:m,k:n) - 2*v*(v'*R(k:m,k:n));

    %Right Hand Side
    if k<n-1
        x = R(k,k+1:n);
        e = zeros(1,length(x));
        e(1) = norm(x);
        if x(1) == 0
            beta = 1;
        else
            beta = sign(x(1));
        end
        u = beta*e + x;
        v = u / norm(u);
        R(k:m,k+1:n) = R(k:m,k+1:n) - 2*R(k:m,k+1:n)*v'*v;
    end
end

R = R(1:n,1:n);

%Clean to use GKbdiag code (does not recognize e_mach as zero)
for i = 1:m
    for j = 1:n
        if abs(R(i,j))<10^-15
            R(i,j)=0;
        end
    end
end
end

function svds = iterGKbdiag(B)
% an illustrative code for computing the singular values of a bidiagonal B
nsvd = 0;
n = min(size(B));
bsize = n;
svds = zeros(n,1);
while nsvd < n
    if bsize >= 2
        B = onestepGKbdiag(B);
        if abs(B(end-1,end)) < eps/2*(abs(B(end-1,end-1)) +
+abs(B(end,end)))
            nsvd = nsvd + 1;
    end
end

```

```

        svals(nsvd) = abs(B(end,end));
        B = B(1:end-1,1:end-1);
        bsize = bsize - 1;
    end
else
    nsvd = nsvd + 1;
    svals(nsvd) = abs(B(1,1));
end
end
end

function [D,CSL,CSR] = onestepGKbdiag(B)
% following Golub & van Loan Alg. 8.6.1:
% apply Givens rotation G' on the left (pre-multiply by [c -s; s c]), and
% apply G on the right (post-multiply by [c s; -s c]). 
n = min(size(B));
if nnz(tril(B,-1)) > 0 || nnz(triu(B,2)) > 0
    error('Input matrix B is not upper bidiagonal');
end
if n >= 3
    T = [B(n-1,n-1)^2+B(n-2,n-1)^2 B(n-1,n-1)*B(n-1,n);
          B(n-1,n-1)*B(n-1,n) B(n,n)^2+B(n-1,n)^2];
else
    T = B'*B;
end
CSL = zeros(2,n-1); CSR = zeros(2,n-1);
evT = eig(T);
[~,idx] = min(abs(evT-(B(n,n)^2+B(n-1,n)^2)));
lambda = evT(idx);
u = [B(1,1)^2-lambda; B(1,1)*B(1,2)];
[c,s] = givens(u);
B(1:2,1:2) = B(1:2,1:2)*[c s; -s c];
CSR(:,1) = [c; s];

for k = 1 : n-2
    [c,s] = givens(B(k:k+1,k));
    B(k:k+1,k:k+2) = [c s; -s c]'*B(k:k+1,k:k+2);
    B(k+1,k) = 0;
    CSL(:,k) = [c; s];

    [c,s] = givens(B(k,k+1:k+2)');
    B(k:k+2,k+1:k+2) = B(k:k+2,k+1:k+2)*[c s; -s c];
    B(k,k+2) = 0;
    CSR(:,k+1) = [c; s];
end
[c,s] = givens(B(n-1:n,n-1));
B(n-1:n,n-1:n) = [c -s; s c]*B(n-1:n,n-1:n);
B(n,n-1) = 0;
D = B;

end

function [c,s] = givens(u)
    if u(2) == 0

```

```
c = 1; s = 0;
else
    if abs(u(2)) > abs(u(1))
        tau = -u(1)/u(2); s = 1/sqrt(1+tau^2); c = s*tau;
    else
        tau = -u(2)/u(1); c = 1/sqrt(1+tau^2); s = c*tau;
    end
end
end
```

Published with MATLAB® R2023b