Midterm

Friday, March 15, 2024 6:01 PM

1. **[Q1]** (a) Let $f(x_1, x_2) = x_1^2 \ln x_2$ where $x_2 > 0$. Find the relative condition number of f. If $x_1 \approx 1$, for what values of x_2 is this evaluation ill-conditioned?

   (b) Let $A \in \mathbb{R}^{n \times n}$ and (λ, v) be an eigenpair such that $Av = \lambda v$. Let $(\hat{\lambda}, \hat{v})$ be a numerically computed approximation to (λ, v). Assume that $\|v\|_2 = \|\hat{v}\|_2 = 1$. Find the vector x, such that $(\hat{\lambda}, \hat{v})$ is an eigenpair of $A + \Delta A$, where $\Delta A = x\hat{v}^H$. Give the expression of $\|\Delta A\|_2$ (should not contain x), and explain why $\|A\hat{v} - \hat{\lambda}\hat{v}\|_2 \leq \mathcal{O}(\|A\|_2 \epsilon_{mach})$ means $(\hat{\lambda}, \hat{v})$ is computed by a backward stable algorithm.

   (c) Consider $I_n := \int_0^1 \frac{x^n}{x+10} dx$, which satisfies $I_n + 10I_{n-1} = \int_0^1 x^{n-1} dx = \frac{1}{n}$. Given $I_0 = \ln \frac{11}{10}$ (with $n = 1$), we can evaluate $I_1 = -10I_0 + \frac{1}{1}$, $I_2 = -10I_1 + \frac{1}{2}$, etc. Use this method to evaluate I_{20}, then call MATLAB's `integral` function with absolute and relative tolerance set to ϵ_{mach} to evaluate I_{20} numerically. Discuss your findings.

a) $f(x_1, x_2) = x_1^2 \ln x_2 \quad x_2 > 0$

$$\kappa_{rel}^2 = \kappa_{abs}^2 \frac{\|x\|_2^2}{\|f(x)\|_2^2} = \frac{\|J_f\|_2^2 \|x\|_2^2}{\|f(x)\|_2^2} = \frac{\|\nabla f\|_2^2 \|x\|_2^2}{\|f(x)\|_2^2}$$

$$= \left(4x_1^2 \ln^2 x_2 + \frac{x_1^4}{x_2^2}\right) \frac{(x_1^2 + x_2^2)}{x_1^4 \ln^2 x_2}$$

$$= 4\frac{(x_1^2 + x_2^2)}{x_1^2} + \frac{x_1^2 + x_2^2}{x_2^2 \ln^2 x_2}$$

$$= 4\left(1 + \left(\frac{x_2}{x_1}\right)^2\right) + \frac{x_1^2 + x_2^2}{x_2^2 \ln^2 x_2}$$

Let $x_1 \approx 1$ $\quad \kappa_{rel}^2 \approx 4(1 + x_2^2) + \frac{1}{x_2^2 \ln^2 x_2} + \frac{1}{\ln^2 x_2}$

$$\lim_{x_2 \to 0} \kappa_{rel}^2 = \infty \quad \lim_{x_2 \to 1^+} \kappa_{rel}^2 = \infty$$

$$\lim_{x_2 \to 1^-} \kappa_{rel}^2 = \infty \quad \lim_{x_2 \to \infty} \kappa_{rel}^2 = \infty$$

b) $(A + \Delta A)\hat{v} = \hat{\lambda}\hat{v} = (A + x\hat{v}^H)\hat{v}$

$= A\hat{v} + x\|\hat{v}\|$

$\Rightarrow x = \hat{\lambda}\hat{v} - A\hat{v}$

Consider $\|\Delta A\|_2 = \sup_{\|u\|_2 = 1} \|\Delta A u\|_2 \geq \|\Delta A \hat{v}\|_2 = \|x\hat{v}^H\hat{v}\|_2 = \|x\|_2 = \|\hat{\lambda}\hat{v} - A\hat{v}\|_2$

Also, $\|\Delta A\|_2 = \|x\hat{v}\|_2 \leq \|x\|_2 \|\hat{v}\|_2 = \|x\|_2 = \|\hat{\lambda}\hat{v} - A\hat{v}\|_2$

It Follows $\|\Delta A\|_2 = \|\hat{\lambda}\hat{v} - A\hat{v}\|_2$

An algorithm is backward stable if

$$\frac{\|f(x) - \hat{f}(x)\|_2}{\|f(x)\|_2} \leq O(\varepsilon_m)$$

Equivalently, for matrices in this case

$$\frac{\|A\hat{v} - \hat{\lambda}\hat{v}\|_2}{\|A\|_2\|\hat{v}\|_2} \leq O(\varepsilon_m) \Rightarrow \|A\hat{v} - \hat{\lambda}\hat{v}\|_2 \leq O(\varepsilon_m \|A\|_2)$$

then this algo is backward stable

c) The numbers are very different which indicates that this is poorly conditioned.

2) The numbers are very different which indicates that this is poorly conditioned.
See attached code for results at the end

2. **[Q2]** (a) Given a 6×4 matrix A with all nonzero entries, illustrate the procedure of Golub-Kahan bidiagonalization, and explain how to compute all singular values of A.

(b) Let $x \in \mathbb{R}^n$, and consider the vector $z = \begin{bmatrix} 0_{n-1} \\ \|x\|_2 \\ x \end{bmatrix} \in \mathbb{R}^{2n}$. Find the Householder reflector $H = I - 2vv^T$ that reduces z such that Hz is a multiple of e_1 (sufficient to find the expression of v). For $y = \begin{bmatrix} 0_n \\ x \end{bmatrix} \in \mathbb{R}^{2n}$, give the simplfied expression of Hy.

(c) Let $U \in \mathbb{R}^{n \times m}$ with $m \ll n$, and $S \in \mathbb{R}^{s \times n}$ with s being a small multiple of m (e.g., $s = 4m$) and $s \ll n$. Suppose that S is well-conditioned (therefore has full rank s), and U is of full rank m. Assume that $SU \in \mathbb{R}^{s \times m}$ has orthonormal columns. Show that $P = I - U(SU)^T S$ is a projector and find its null space and the range. (hint: consider the reduced QR factorization of S^T)

a) Let X be an undetermined coefficient in $\mathbb{F}$.
Let $A \in \mathbb{F}^{6\times4}$. It follows

$$A = \begin{bmatrix} X & X & X & X \\ X & X & X & X \\ X & X & X & X \\ X & X & X & X \\ X & X & X & X \\ X & X & X & X \end{bmatrix} \quad U_1 A = \begin{bmatrix} X & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \end{bmatrix}$$

Construct H_{R_1} $V_1 = \begin{bmatrix} 1 & 0 \\ 0 & H_{R_1} \end{bmatrix}$

construct $H_{L_1} = U_1$

$$U_1 A V_1 = \begin{bmatrix} X & X & 0 & 0 \\ 0 & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \\ 0 & X & X & X \end{bmatrix} \quad U_2 U_1 A V_1 = \begin{bmatrix} X & X & 0 & 0 \\ 0 & X & X & X \\ 0 & 0 & X & X \\ 0 & 0 & X & X \\ 0 & 0 & X & X \\ 0 & 0 & X & X \end{bmatrix}$$

Construct H_{R_2} $V_2 = \begin{bmatrix} I_2 & 0 \\ 0 & H_{R_2} \end{bmatrix}$

construct H_{L_2} $U_2 = \begin{bmatrix} 1 & 0 \\ 0 & H_{L_2} \end{bmatrix}$

$$U_2 U_1 A V_1 V_2 = \begin{bmatrix} X & X & 0 & 0 \\ 0 & X & X & 0 \\ 0 & 0 & X & X \\ 0 & 0 & X & X \\ 0 & 0 & X & X \\ 0 & 0 & X & X \end{bmatrix} \quad U_3 U_2 U_1 A V_1 V_2 = \begin{bmatrix} X & X & 0 & 0 \\ 0 & X & X & 0 \\ 0 & 0 & X & X \\ 0 & 0 & 0 & X \\ 0 & 0 & 0 & X \\ 0 & 0 & 0 & X \end{bmatrix}$$

Construct H_{L_3} $U_3 = \begin{bmatrix} I_2 & 0 \\ 0 & H_{L_3} \end{bmatrix}$

Construct

Construct H_{L3} $U_3 = \begin{bmatrix} I_2 & 0 \\ 0 & H_{L3} \end{bmatrix}$

Construct H_{L4} $U_4 = \begin{bmatrix} I_3 & 0 \\ 0 & H_{L4} \end{bmatrix}$

$$B = U_4 U_3 U_2 U_1 A V_1 V_2 = \begin{bmatrix} X & X & 0 & 0 \\ 0 & X & X & 0 \\ 0 & 0 & X & X \\ 0 & 0 & 0 & X \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

2nd Phase Compute trailing 2x2 submatrix T of B^TB

Find $\lambda(T)$ that is closest to (2,2) element of T form $u = \begin{bmatrix} (1,1) \text{ element of } B^TB - \lambda \\ (2,1) \text{ element of } B^TB \end{bmatrix}$

Determine $G_1 \ni G_1^T u = \begin{bmatrix} * \\ 0 \end{bmatrix}$

Then update $B = B\begin{bmatrix} G_1 & 0 \\ 0 & I \end{bmatrix}$

After 2-3 iterations at each step k, the k^{th} last diagonal of B^TB converges to σ_k

b) $v = z - \|z\|_2 e_1$ $\quad \|z\|_2^2 = 2\|x\|_2^2 \Rightarrow \|z\|_2 = \sqrt{2}\,\|x\|_2$

$$v = \begin{bmatrix} -\sqrt{2}\|x\|_2 \\ 0_{n-2} \\ \|x\|_2 \\ x \end{bmatrix}$$

$$v^Tv = 2\|x\|_2^2 + \|x\|_2^2 + \|x\|_2^2 = 4\|x\|_2^2$$

$$vv^T = \begin{bmatrix} 2\|x\|_2^2 & 0 & -\sqrt{2}\|x\|_2^2 & -\sqrt{2}\|x\|_2 x^T \\ 0 & 0 & 0 & 0 \\ -\sqrt{2}\|x\|_2^2 & 0 & \|x\|_2^2 & \|x\|_2 x^T \\ -\sqrt{2}\|x\|_2 x & 0 & \|x\|_2 x & xx^T \end{bmatrix}$$

$$H = I - 2\frac{vv^T}{v^Tv} = I - \frac{1}{2\|x\|_2^2} vv^T$$

$$Hz = \begin{bmatrix} 0 \\ 0 \\ \|x\|_2 \\ x \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 2 & 0 & \sqrt{2} & -\sqrt{2}\frac{x^T}{\|x\|_2} \\ 0 & 0 & 0 & 0 \\ \sqrt{2} & 0 & 1 & \frac{x^T}{\|x\|_2} \\ -\sqrt{2}\frac{x}{\|x\|_2} & 0 & \frac{x}{\|x\|_2} & \frac{xx^T}{\|x\|_2^2} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ \|x\|_2 \\ x \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ \|x\|_2 \\ x \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 0 + 0 - \sqrt{2}\|x\|_2 - \sqrt{2}\|x\|_2 \\ 0 + 0 + 0 + 0 \\ 0 + 0 + \|x\|_2 + \|x\|_2 \\ 0 + 0 + x + x \end{bmatrix} = \begin{bmatrix} \sqrt{2}\|x\|_2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = w \checkmark$$

$$Hy = \begin{bmatrix} 0 \\ 0 \\ 0 \\ x \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 2 & 0 & -\sqrt{2} & -\sqrt{2}\frac{x^T}{\|x\|_2} \\ 0 & 0 & 0 & 0 \\ -\sqrt{2} & 0 & 1 & \frac{x^T}{\|x\|_2} \\ -\sqrt{2}\frac{x}{\|x\|_2} & 0 & \frac{x}{\|x\|_2} & \frac{xx^T}{\|x\|_2^2} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 0 \\ x \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \\ x \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -\sqrt{2}\|x\|_2 \\ 0 \\ \|x\|_2 \\ x \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2}\|x\|_2 \\ 0 \\ -\frac{\|x\|_2}{2} \\ \frac{1}{2}x \end{bmatrix}$$

c) $P^2 = (I - U(SU)^TS)(I - U(SU)^TS) = I - 2U(SU)^TS + U\cancel{(SU)^TSU}(SU)^TS$
$= I - U(SU)^TS = P$

Thus P is a projector.

$I - P = I - I + U(SU)^TS = U(SU)^TS$

$R(P) = N(I-P) = N(U(SU)^TS) = R(S)^\perp = N(S)$

$N(P) = R(I-P) = R(U(SU)^TS) = R(U)$

$I - P = I - I + U(SU)^T S = U(SU)^T S$

$R(P) = N(I-P) = N(U(SU)^T S) = R(S)^\perp = N(S)$

$N(P) = R(I-P) = R(U(SU)^T S) = R(U)$

3. **[Q3]** (a) For $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$, let $A^\dagger = (A^T A)^{-1} A^T$. Show that $\|A^\dagger\|_2 = \frac{1}{\sigma_n(A)}$. (assume that A has full column rank)

(b) Let $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$, where $A_1 \in \mathbb{R}^{n\times n}$ is nonsingular. Show that $\sigma_n(A) \geq \sigma_n(A_1)$ (explore the relation between $\frac{\|Ax\|_2}{\|x\|_2}$ and $\frac{\|A_1 x\|_2}{\|x\|_2}$), and $\|A^\dagger\|_2 \leq \|A_1^{-1}\|_2$.

(c) Define the numerical rank of $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$ as $\text{rank}(A,\epsilon) = \max\{k : \sigma_k \geq \epsilon\}$ $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n)$. If A has numerical rank $k < n$ for a given ϵ, find a numerically full rank B satisfying $\inf_{\text{rank}(B,\epsilon)=n} \|A - B\|_F$ and show that $\|B - A\|_F \leq \sqrt{n-k}\,\epsilon$.

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

$= V\Sigma^{-2} \cancel{V^T V} \Sigma U^T$

$= V\Sigma^{-1} U^T$

$\|A^\dagger\|_2 = \sqrt{\rho(A^\dagger)} = \sqrt{\rho(V\Sigma^{-1}U^T)} = \sqrt{\rho(\Sigma^{-1})}$

Since $\sigma_n(A) \leq \sigma_{i \leq n}(A)$, $\frac{1}{\sigma_{i\leq n}(A)} \leq \frac{1}{\sigma_n(A)}$.

Thus $\sqrt{\rho(\Sigma^{-1})} = \frac{1}{\sigma_n(A)}$

b) $\sigma_{min}^2(A) = \inf_{\|x\|_2=1} \left\|\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} x\right\|_2^2 = \inf_{\|x\|_2=1}\left(\|A_1 x\|_2^2 + \|A_2 x\|_2^2\right) \geq \inf_{\|x\|_2=1} \|A_1 x\|_2^2 = \sigma_{min}^2(A_1)$

$A_1^\dagger = (A_1^T A_1)^{-1} A_1^T = A_1^{-1}$

$\sigma_n(A) \geq \sigma_n(A_1) \Rightarrow \frac{1}{\sigma_n(A_1)} \geq \frac{1}{\sigma_n(A)} \Rightarrow \|A_1^\dagger\|_2 \geq \|A^\dagger\|_2$

$\Rightarrow \|A_1^{-1}\|_2 \geq \|A^\dagger\|_2$

c) Let $A = U\Sigma_A V^T$.

Define $B = U\Sigma_B V^T \ni \Sigma_{B_{ii}} = \begin{cases} \sigma_i & 1 \leq i \leq k \\ \epsilon & k+1 < i < n \end{cases}$

Assume $\exists\, C \ni \text{rank}(C,\epsilon) = n$ and $\|A - C\|_F \leq \|A - B\|_F$

$\sum_{i=1}^n (\sigma_i - c_i)^2 = \sum_{i=1}^n \sigma_i^2 - 2\sigma_i c_i + c_i^2$

$\geq \sum_{i=1}^n \sigma_i^2 - 2\sigma_i c_i + \epsilon^2$ $\qquad c_i^2 \geq \epsilon^2$ adding less

$\geq \sum_{i=k+1}^n \sigma_i^2 + \epsilon^2 + \sum_{i=1}^n 2\sigma_i(-c_i)$

$\geq \sum_{i=k+1}^n \sigma_i^2 + \epsilon^2 + 2\sum_{i=1}^n 2\sigma_i(-\epsilon)$ $\qquad -c_i < -\epsilon$ (subtracting less)

$\geq \sum_{i=k+1}^n \sigma_i^2 + \epsilon^2 - 2\sigma_i\epsilon$ $\qquad$ subtracting less

$= \sum_{i=k+1}^n (\sigma_i - \epsilon)^2$

$= \|\Sigma_A - \Sigma_B\|_F$

$= \|U(\Sigma_A - \Sigma_B)V^T\|_F$

$= \|U\Sigma_A V^T - U\Sigma_B V^T\|_F$

$= \|A - B\|_F$

Thus B reaches the infimum.

It follows, $\|B - A\|_F^2 = \|U\Sigma_B V^T - U\Sigma_A V^T\|_F^2$

$= \|U(\Sigma_B - \Sigma_A)V^T\|_F^2$

It follows, $\|B-A\|_F^2 = \|U\Sigma_B V^T - U\Sigma_A V^T\|_F^2$

$$= \|U(\Sigma_B - \Sigma_A)V^T\|_F^2$$
$$= \|\Sigma_B - \Sigma_A\|_F^2$$
$$= \sum_{i=k+1}^{n} (\Sigma - \sigma_i)^2$$
$$\leq \sum_{i=k+1}^{n} \Sigma^2$$
$$= (n-k)\Sigma^2$$
$$\Rightarrow \|B-A\|_F \leq \sqrt{n-k}\,\Sigma$$

```
i0 = log(11/10);
for i = 1:20
    i0 = -10*i0 + 1/i;
end

fun = @(x) x.^20./(x+10);

i20 = integral(fun,0,1);

i0
i20
```

```
i0 =

   7.4835e+03


i20 =

    0.0043
```

Published with MATLAB® R2023b