2021 Spring MAS 365: Homework 4

posted on Apr 1; due by Apr 8

1. [5+5+10 points] Let A be a nonsingular $n \times n$ matrix, $||\cdot||$ be any natural norm, and $K_p(A) = ||A||_p||A^{-1}||_p$. Let λ_1 be the smallest and λ_n be the largest eigenvalues of the matrix A^tA.

   (a) Show that if λ is an eigenvalue of A^tA, then $0 < \lambda \leq ||A^t||\ ||A||$.

   (b) Show that $K_2(A) = \sqrt{\frac{\lambda_n}{\lambda_1}}$. (Hint: $||A|| = \sqrt{\rho(A^tA)} = \sqrt{\rho(AA^t)}$.)

   (c) Show that $K_2(A) \leq \sqrt{K_1(A)K_\infty(A)}$. (Hint: $||A||_1 = \max_{1\leq j\leq n}\sum_{i=1}^n |a_{ij}|$ and $||A||_\infty = \max_{1\leq i\leq n}\sum_{j=1}^n |a_{ij}|$.)

   Solution:

   (a) Since A is nonsingular, A^tA is nonsingular and this implies $\lambda \neq 0$. Since $A^tA\boldsymbol{x} = \lambda\boldsymbol{x}$,

   $$\begin{aligned}\lambda||\boldsymbol{x}||^2 &= \boldsymbol{x}^tA^tA\boldsymbol{x} = ||A\boldsymbol{x}||^2 \geq 0,\\ |\lambda|\ ||\boldsymbol{x}|| &= ||A^tA\boldsymbol{x}|| \leq ||A^t||\ ||A\boldsymbol{x}|| \leq ||A^t||\ ||A||\ ||\boldsymbol{x}||.\end{aligned}$$

   Dividing both inequalities by $||\boldsymbol{x}||^2$ and $||\boldsymbol{x}||$ for nonzero $\boldsymbol{x}$, we have $0 < \lambda \leq ||A^t||\ ||A||$.

   (b) From (a), λ_1 and λ_n are positive. If λ is an eigenvalue of A^tA, then $\frac{1}{\lambda}$ is an eigenvalue of $(A^tA)^{-1}$. Using the hint, we have $||A||_2 = \sqrt{\rho(A^tA)} = \sqrt{\lambda_n}$ and $||A^{-1}|| = \sqrt{\rho(A^{-1}(A^{-1})^t)} = \sqrt{\rho((A^tA)^{-1})} = \frac{1}{\sqrt{\lambda_1}}$. Therefore, $K_2(A) = \sqrt{\frac{\lambda_n}{\lambda_1}}$.

   (c) The hint implies that $||A||_\infty = ||A^t||_1$ and $||A^{-1}||_\infty = ||(A^{-1})^t||_1$. Using (a), we have

   $$\lambda_n \leq ||A^t||_1\ ||A||_1, \quad \lambda_1^{-1} \leq ||(A^{-1})^t||_1\ ||A^{-1}||_1$$

   Therefore, using (b),

   $$\begin{aligned}K_2(A) = \sqrt{\frac{\lambda_n}{\lambda_1}} &\leq \sqrt{||A^t||_1||A||_1||(A^{-1})^t||_1||A^{-1}||_1}\\ &= \sqrt{||A||_\infty||A||_1||A^{-1}||_\infty||A^{-1}||_1} = \sqrt{K_1(A)K_\infty(A)}\end{aligned}$$

2. [10+10+10 points] Prove Theorem 7.33 in the textbook using mathematical induction as follows:

   (a) Show that $\langle \boldsymbol{r}^{(1)}, \boldsymbol{v}^{(1)}\rangle = 0$.

   (b) Assume that $\langle \boldsymbol{r}^{(k)}, \boldsymbol{v}^{(j)}\rangle = 0$, for each $k \leq l$ and $j = 1, 2, \ldots, k$, and show that this implies that $\langle \boldsymbol{r}^{(l+1)}, \boldsymbol{v}^{(j)}\rangle = 0$, for each $j = 1, 2, \ldots, l$.

   (c) Show that $\langle \boldsymbol{r}^{(l+1)}, \boldsymbol{v}^{(l+1)}\rangle = 0$.

   Solution:

(a) Since

$$\boldsymbol{r}^{(1)} = \boldsymbol{b} - A\boldsymbol{x}^{(1)} = \boldsymbol{b} - A\left(\boldsymbol{x}^{(0)} + \frac{\langle \boldsymbol{v}^{(1)},\, \boldsymbol{r}^{(0)}\rangle}{\langle \boldsymbol{v}^{(1)},\, A\boldsymbol{v}^{(1)}\rangle}\boldsymbol{v}^{(1)}\right) = \boldsymbol{r}^{(0)} - \frac{\langle \boldsymbol{v}^{(1)},\, \boldsymbol{r}^{(0)}\rangle}{\langle \boldsymbol{v}^{(1)},\, A\boldsymbol{v}^{(1)}\rangle} A\boldsymbol{v}^{(1)},$$

we have that

$$\langle \boldsymbol{r}^{(1)},\, \boldsymbol{v}^{(1)}\rangle = \langle \boldsymbol{r}^{(0)},\, \boldsymbol{v}^{(1)}\rangle - \frac{\langle \boldsymbol{v}^{(1)},\, \boldsymbol{r}^{(0)}\rangle}{\langle \boldsymbol{v}^{(1)},\, A\boldsymbol{v}^{(1)}\rangle} \langle A\boldsymbol{v}^{(1)},\, \boldsymbol{v}^{(1)}\rangle = 0$$

(b) Since

$$\boldsymbol{r}^{(l+1)} = \boldsymbol{r}^{(l)} - \frac{\langle \boldsymbol{v}^{(l+1)},\, \boldsymbol{r}^{(l)}\rangle}{\langle \boldsymbol{v}^{(l+1)},\, A\boldsymbol{v}^{(l+1)}\rangle} A\boldsymbol{v}^{(l+1)},$$

we have that

$$\langle \boldsymbol{r}^{(l+1)},\, \boldsymbol{v}^{(j)}\rangle = \langle \boldsymbol{r}^{(l)},\, \boldsymbol{v}^{(j)}\rangle - \frac{\langle \boldsymbol{v}^{(l+1)},\, \boldsymbol{r}^{(l)}\rangle}{\langle \boldsymbol{v}^{(l+1)},\, A\boldsymbol{v}^{(l+1)}\rangle} \langle A\boldsymbol{v}^{(l+1)},\, \boldsymbol{v}^{(j)}\rangle$$

for all $j = 1, \ldots, l$. The first term of the right-hand side is zero by the assumption, and the second term is also zero due to the A-orthogonality of $\{\boldsymbol{v}^{(1)}, \ldots, \boldsymbol{v}^{(l+1)}\}$. Therefore, $\langle \boldsymbol{r}^{(l+1)},\, \boldsymbol{v}^{(j)}\rangle = 0$ for all $j = 1, \ldots, l$.

(c) Similarly, we have

$$\langle \boldsymbol{r}^{(l+1)},\, \boldsymbol{v}^{(l+1)}\rangle = \langle \boldsymbol{r}^{(l)},\, \boldsymbol{v}^{(l+1)}\rangle - \frac{\langle \boldsymbol{v}^{(l+1)},\, \boldsymbol{r}^{(l)}\rangle}{\langle \boldsymbol{v}^{(l+1)},\, A\boldsymbol{v}^{(l+1)}\rangle} \langle A\boldsymbol{v}^{(l+1)},\, \boldsymbol{v}^{(l+1)}\rangle = 0.$$

3. [10+10 points]

   (a) Implement the SOR method with the optimal choice of w for the positive definite and tridiagonal matrix via MATLAB grader.

   (b) Implement the steepest descent method via MATLAB grader.

Solution:

(a)
```
function [ws xs Ns] = SOR(A, b, x0, epsilon, N)
    Tj = inv(diag(diag(A)))*(-diag(diag(A,1),1)-diag(diag(A,-1),-1));
    rhoTj = max(abs(eig(Tj)));
    ws = 2/(1+sqrt(1-rhoTj^2));
    n = size(A,1);
    xs = x0;
    for k=1:N
        xs_prev = xs;
        for i=1:n
            xs(i) = (1-ws)*xs(i) + ws/A(i,i)*(-A(i,[1:i-1 i+1:n])*xs([1:i-1 i+1:n]) ...
                    + b(i));
        end
        if (norm(xs - xs_prev, Inf)/norm(xs, Inf) < epsilon)
            break;
        end
    end
    Ns = k;
end
```

(b)
```
function [xs Ns] = steepest_descent(A, b, x0, epsilon, N)
    x = x0;
    for k=1:N
        xprev = x;
        r = b - A*x;
        t = r'*r/(r'*A*r);
        x = x + t*r;
        if (norm(x - xprev, Inf)/norm(x,Inf) < epsilon)
            break;
        end
    end
    xs = x;
    Ns = k;
end
```