HINTS:

PROBLEM 1: Note that $\| A - c_j s r_i \|_2 = \| (u_1 \sqrt{\varepsilon} v_1^T + \varepsilon U_2 V_2^T) - c_j s r_i \|_2$
$\geq \| u_1 \sqrt{\varepsilon} v_1^T - c_j s r_i \|_2 - \| \varepsilon U_2 V_2^T \|_2$
Prove that $\forall s \in \mathbb{R}$ we have $\| u_1 \sqrt{\varepsilon} v_1^T - c_j s r_i \|_2 \geq \sqrt{\frac{\varepsilon}{2}}$:
to do so, multiply by U^T on the left, by V on the right, then use that $\| [\underline{a} \; B] \|_2 \geq \| \underline{a} \|_2$ for a column vector $\underline{a}$.

PROBLEM 3: To prove that for a given $x \in \mathbb{R}^{n-1} \setminus \{0\}$ you have $x^T S x > 0$, look for a vector $\tilde{x} = \begin{bmatrix} ? \\ x \end{bmatrix} \in \mathbb{R}^n$ such that $x^T S x = \tilde{x}^T A \tilde{x}$.

PROBLEM 4:

(1) where is the maximum element of a symm. pos. def. matrix?

(2) they correspond to P_______L submatrices

(3) Look at the entry corresponding to index sets I and J of cardinality k
- How do you obtain the submatrix $(XY)[I,J]$ as the product of a submatrix of X and a submatrix Y?
- Apply Cauchy-Binet to this product
- Relate the elements appearing in Cauchy-Binet to elements of $C_k(X)$ and $C_k(Y)$

(4) Apply point (3) to $A = LL^T$, noting that $C_k(L^T) = C_k(L)^T$

(5) Note that the statement can be rephrased as "Among the elements of $C_k(A)$, the max absolute value is attained on the diagonal".

SOLUTIONS

PROBLEM 1

Note that $A = \sqrt{\varepsilon} \cdot u_1 v_1^T + \varepsilon U_2 V_2^T$ and $\| U_2 V_2^T \|_2 \leq 1$, so

$$\| A - c_j s r_i \|_2 \geq \| \sqrt{\varepsilon} u_1 v_1^T - c_j s r_i \|_2 - \| \varepsilon U_2 V_2^T \|_2 \geq \| \sqrt{\varepsilon} u_1 v_1^T - c_j s r_i \|_2 - \varepsilon.$$

Therefore, it is sufficient to prove that $\| \sqrt{\varepsilon} u_1 v_1^T - c_j s r_i \|_2 \geq \sqrt{\frac{\varepsilon}{2}} \quad \forall s \in \mathbb{R}$

$$\| \sqrt{\varepsilon} u_1 v_1^T - c_j s r_i \|_2 \underset{\uparrow \; U, V \text{ orthogonal}}{=} \| \sqrt{\varepsilon} \underbrace{U^T u_1}_{e_1} \underbrace{v_1^T V}_{e_1^T} - U^T c_j s r_i V \|_2 =$$

$$= \left\| \begin{bmatrix} \sqrt{\varepsilon} & 0 \\ 0 & O \end{bmatrix} - s \cdot \begin{bmatrix} u_1^T c_j \\ U_2^T c_j \end{bmatrix} \begin{bmatrix} r_i v_1 & r_i V_2 \end{bmatrix} \right\|_2 \geq$$

for a vector a and matrix B, we have $\| [a \; B] \|_2 \geq \| a \|_2$

$$\geq \left\| \begin{bmatrix} \sqrt{\varepsilon} - s \cdot u_1^T c_j \cdot r_i v_1 \\ - s \cdot U_2^T c_j \cdot r_i v_1 \end{bmatrix} \right\|_2 =$$

- $u_1^T c_j = u_1^T \cdot U \Sigma V^T e_j = e_1^T \Sigma V e_j = \sqrt{\varepsilon}\, e_1^T V^T e_j = \sqrt{\varepsilon}\, v_1^T e_j = \sqrt{\frac{\varepsilon}{n}}$
- for $k = 2, 3, \ldots, n$ we have $u_k^T c_j = u_k^T U \Sigma V^T e_j = \varepsilon e_k^T V^T e_j = \varepsilon v_{jk}$ $\Rightarrow \| U_2^T c_j \|_2^2 = \varepsilon^2 \| V(j, 2:n) \|_2^2 = \varepsilon^2 \left(1 - \frac{1}{n}\right)$

because $\| V(j,:) \| = 1$ by orthogonality, and $V_{j1} = 1/\sqrt{n}$

$$= \left(\left(\sqrt{\varepsilon} - s \sqrt{\frac{\varepsilon}{n}} r_i v_1\right)^2 + s^2 \cdot (r_i v_1)^2 \cdot \varepsilon^2 \left(1 - \frac{1}{n}\right) \right)^{1/2}$$

let $g := s r_i v_1$

$$= \left(\varepsilon + g^2 \frac{\varepsilon}{n} - 2 \frac{\varepsilon}{\sqrt{n}} g + g^2 \varepsilon^2 \left(1 - \frac{1}{n}\right) \right)^{1/2} =$$

$$= \sqrt{\varepsilon} \cdot \left(g^2 \underbrace{\left(\frac{1}{n} + \varepsilon - \frac{\varepsilon}{n} \right)}_{= \frac{1}{n} + \frac{1}{n-1} - \frac{1}{n(n-1)} = \frac{2}{n}} - \frac{2}{\sqrt{n}} g + 1 \right)^{1/2} = \sqrt{\varepsilon} \left(\underbrace{\frac{2}{n} g^2 - \frac{2}{\sqrt{n}} g + 1}_{\text{quadratic function of } g} \right)^{1/2} \geq$$

quadratic function of g, has its min in $g = \frac{2/\sqrt{n}}{2 \cdot (2/n)} = \frac{\sqrt{n}}{2}$

$$\geq \sqrt{\varepsilon} \cdot \left(\frac{2}{n} \cdot \frac{n}{4} - \frac{2}{\sqrt{n}} \frac{\sqrt{n}}{2} + 1 \right)^{1/2} = \sqrt{\frac{\varepsilon}{2}}$$

This concludes the proof. ∎

PROBLEM 3

To prove that S is SPD, it is sufficient to prove that for all $x \in \mathbb{R}^{n-1} \setminus \{0\}$ we have $x^T S x > 0$. If we find a vector $\tilde{x} \in \mathbb{R}^n$ of the form $\tilde{x} = \begin{bmatrix} y \\ x \end{bmatrix}$ such that $x^T S x = \tilde{x}^T A \tilde{x}$ then we are done because $\tilde{x}^T A \tilde{x} > 0$ by positive definiteness of A.

We have that, taking $y := -\frac{1}{\alpha} b^T x$,

$$\tilde{x}^T A \tilde{x} = \begin{bmatrix} -\frac{1}{\alpha} b^T x & x^T \end{bmatrix} \begin{bmatrix} \alpha & b^T \\ b & C \end{bmatrix} \begin{bmatrix} -\frac{1}{\alpha} b^T x \\ x \end{bmatrix} = \frac{1}{\alpha}(b^T x)^2 - \frac{2}{\alpha}(b^T x)^2 + x^T C x =$$

$$= x^T C x - \frac{1}{\alpha}(b^T x)^2 = x^T \left(C - \frac{1}{\alpha} b b^T \right) x = x^T S x. \quad \blacksquare$$

PROBLEM 4

1) For $|I| = 1$, $|\det(A(I,I))| = |a_{ii}|$; SPD matrices have (at least one of) their elements of maximum absolute value on the diagonal
2) Principal submatrices (the ones for which $I = J$)
3) Note that, for $|I| = |J| = k$,

$$(XY)[I,J] = \underbrace{X(I,:)}_{k \times n \text{ matrix}} \cdot \underbrace{Y(:,J)}_{n \times k \text{ matrix}}$$

By Cauchy-Binet formula,

$$C_k(XY)_{IJ} \overset{\text{def.}}{=} \det\big((XY)(I,J)\big) = \det\big(X(I,:)\,Y(:,J)\big) = \sum_{|K|=k} \det(X(I,K)) \det(Y(K,J))$$

$$\overset{\text{def.}}{=} \sum_{|K|=k} (C_k(X))_{IK} (C_k(Y))_{KJ} =$$

$$= \underbrace{(C_k(X))(I,:)}_{I\text{-th row of } C_k(X)} \cdot \underbrace{(C_k(Y))(:,J)}_{J\text{-th column of } C_k(Y)}$$

This proves that $C_k(XY) = C_k(X) \cdot C_k(Y)$

4) Consider a Cholesky decomposition $A = LL^T$

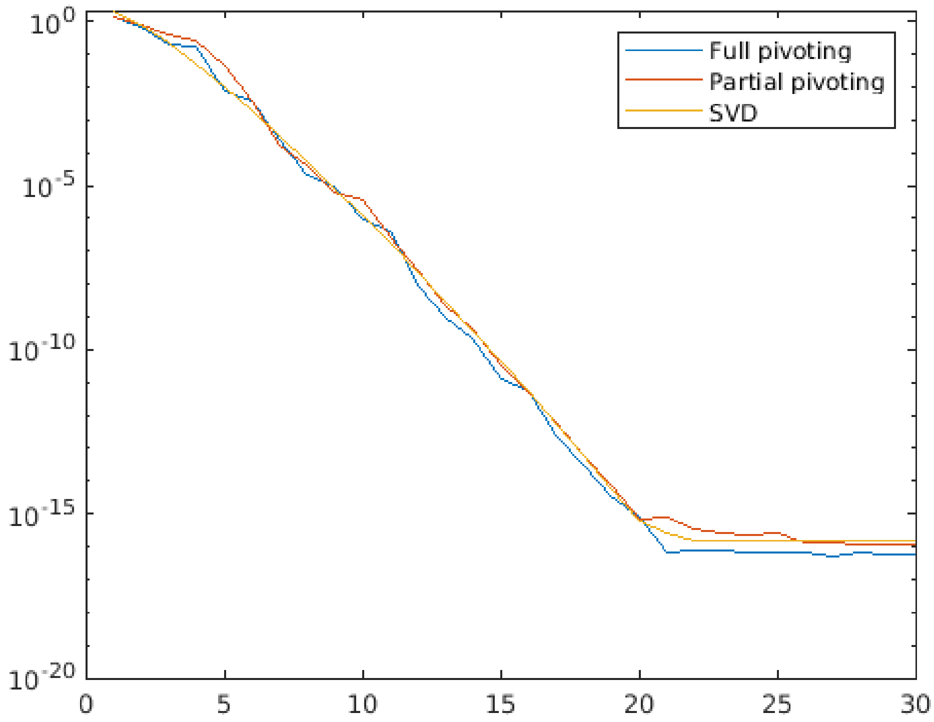
We have that $C_k(A) = C_k(LL^T) = C_k(L)\, C_k(L^T) \overset{\text{follows from the def. of } C_k(\cdot)}{=} \underbrace{C_k(L)}_{\text{triangular, diagonal elements} \neq 0} \cdot C_k(L)^T$

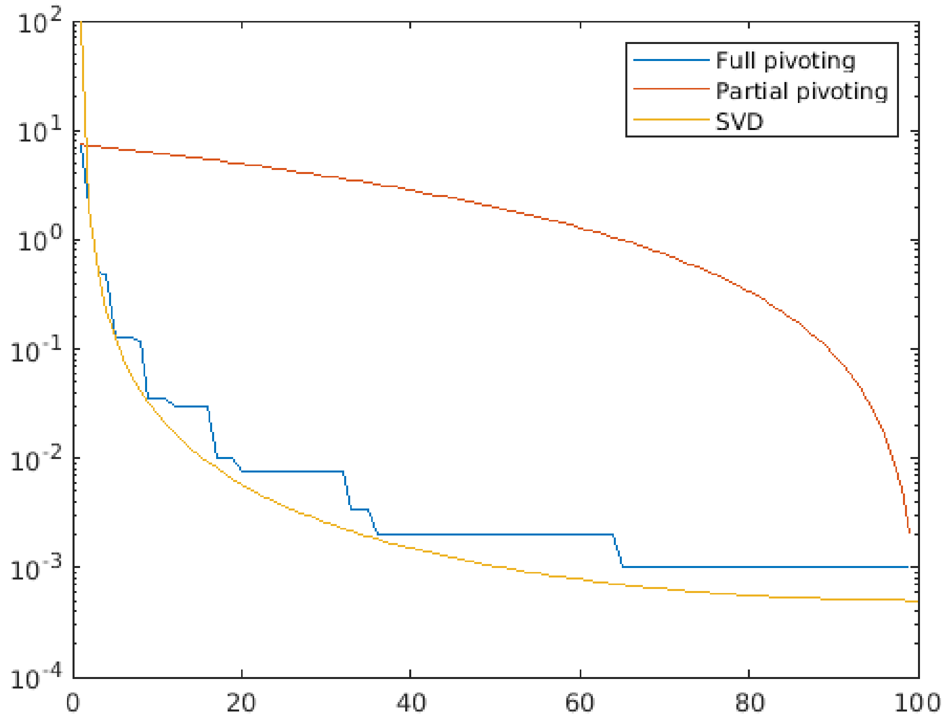
therefore $C_k(A)$ is SPD.

5) The max of the first line of (1) is the max absolute value of an entry of $C_k(A)$; as $C_k(A)$ is SPD (one of) its max elements is on the diagonal and therefore it is the determinant of a principal submatrix. $\blacksquare$

```
n = 100;
A = hilb(n);
R = 30;

figure(1)
aca(A,R)
A = zeros(n,n);
gamma = 0.1;
for i = 1:n
    for j = 1:n
        A(i,j) = exp(-gamma*abs(i-j)/n);
    end
end

R = 100;
figure(2)
aca(A,R)


function aca(A,R)
    [U,S,V] = svd(A);
    s = diag(S);
    errs1 = zeros(R,1);
    errs2 = zeros(R,1);
    for r = 1:R
        A1 = aca_full(A,r);
        errs1(r) = norm(A1);
        A2 = aca_partial(A,r);
        errs2(r) = norm(A2);
    end
    semilogy(1:R,errs1)
    hold on
    semilogy(1:R,errs2)
    semilogy(1:R,s(1:R))
    legend("Full pivoting","Partial pivoting","SVD")
    hold off
end


function R = aca_full(A,K)
    R = A;
    I = [];
    J = [];
    for r = 1:K
        [~,indmax] = max(abs(R(:)));
        [imax,jmax] = ind2sub(size(R),indmax);
        I = [I;imax];
        J = [J;jmax];
        Delta = R(imax,jmax);
        u = R(:,jmax);
        v = R(imax,:)'/ Delta;
        R = R-u*v';
```

```
    end
end

function R = aca_partial(A,K)
    [m,n] = size(A);
    R = A;
    I = [];
    J = [];
    imax = 1;
    for k = 1:K
        [~,jmax] = max(abs(R(imax,:)));
        Delta = R(imax,jmax);
        if Delta == 0
            if length(I) == min(m,n)-1
                return
            end
        else
            u = R(:,jmax);
            v = R(imax,:)'/ Delta ;
            R = R-u*v';
        end
        I = [I;imax];
        J = [J;jmax];
        w = u;
        w(I) = 0;
        [~,imax] = max(abs(w));
    end
end
```


Full pivoting
Partial pivoting
SVD
10^0
10^{-5}
10^{-10}
10^{-15}
10^{-20}
0
5
10
15
20
25
30


Full pivoting
Partial pivoting
SVD
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
10^{-4}
0
20
40
60
80
100