## Hints

① Write $\det(A^TA)$ in terms of the eigenvalues of $A^TA$ and recall the relation between these and the singular values of $A$.

② 1) Take $B = \text{ones}(n-r, r)$

2) In which point are we using that $U$ has orthonormal columns in the proof of the lemma? Note that for a general matrix $U$, if $\tilde{U} := U \cdot U(I,:)^{-1}$ then $U(I,:)^{-1} = U^+ \tilde{U}$.

pseudo-inverse

③ Induction on $n$

④
- What special form does $\tilde{U}$ have? (what is the submatrix corresponding to the rows in $I$?)
- Prove that $\text{Vol}(\tilde{U}(J,:)) = \tilde{u}_{i^*j^*} \cdot \text{Vol}(\tilde{U}(I,:))$
- Relate $\text{Vol}(\tilde{U}(J,:))$ to $\text{Vol}(U(J,:))$ (and same with $I$)

⑤ Prove that $\|Pu\|_2 = \|(I-P)w\|_2$.
Be careful: what happens if $(I-P)u = 0$?

## Solutions

### Problem 1

Let $A \in \mathbb{R}^{n\times n}$ with singular values $\sigma_1, \dots, \sigma_n$.

We have $\mathrm{Vol}(A) \overset{\text{def.}}{=} |\det(A)| = \sqrt{\det(A)^2} = \sqrt{\det(A^TA)} \overset{\text{relation between det \& eigenvalues}}{=}$

$= \sqrt{\lambda_1(A^TA)\cdots\lambda_n(A^TA)} \overset{A^TA \text{ is pos. definite}}{=} \sqrt{\sigma_1^2\sigma_2^2\cdots\sigma_n^2} \overset{\text{Binet theorem}}{=} \sigma_1\sigma_2\cdots\sigma_n.$ □

### Problem 2

① Let $\tilde{U} := \begin{bmatrix} I_r \\ \mathrm{ones}(n-r, r) \end{bmatrix}$. In this case, $\|\mathrm{ones}(n-r, r)\|_2 =$

$= \|\mathrm{ones}(n-r, 1)\cdot \mathrm{ones}(1, r)\|_2 = \|\mathrm{ones}(n-r, 1)\|_2 \cdot \|\mathrm{ones}(1, r)\|_2 = \sqrt{r(n-r)}$

and $\|\tilde{U}\|_2 = \sqrt{1+r(n-r)}$. Now take an economy QR factorization $[U, R] = qr(\tilde{U})$ and note that the indices corresponding to a max. vol. submatrix in $U$ are the same as $\tilde{U}$. This means that, as $I = \{1, \dots, r\}$ gives a max. vol. submatrix in $\tilde{U}$, $U$ is an example for which the lemma is tight. □

② Let $U \in \mathbb{R}^{n\times r}$, consider index set $I$ corresponding to a maximum volume submatrix; define $\tilde{U} := U \cdot U(I,:)^{-1}$. As in the proof seen in the lecture, $\|\tilde{U}\|_2 \le \sqrt{1+r(n-r)}$.

We have that

$$\frac{1}{\sigma_{\min}(U(I,:))} = \|U(I,:)^{-1}\|_2 = \|U^{\dagger}\tilde{U}\|_2 \le \|U^{\dagger}\|_2 \cdot \|\tilde{U}\|_2$$

($U^{\dagger}$: pseudo-inverse)

$$\le \|U^{\dagger}\|_2 \cdot \sqrt{1+r(n-r)} = \frac{1}{\sigma_{\min}(U)} \cdot \sqrt{1+r(n-r)}.$$ □

If $A = U\Sigma V^T$ is an SVD of $A$, the pseudoinverse is given by $A^{\dagger} = V\Sigma^{\dagger}U^T$, where $(\Sigma^{\dagger})_{ii} = \begin{cases} 1/\sigma_i & \text{if } \sigma_i \neq 0 \\ 0 & \text{otherwise} \end{cases}$

### Problem 3

We prove it by induction. The base case $n=1$ is trivial.

Assume the theorem holds for $n-1$.

We have that

$$Z_n = \left[\begin{array}{c|c} Z_{n-1} & \begin{matrix}-1\\ \vdots \\ -1\end{matrix} \\ \hline \begin{matrix}0 & \cdots & 0\end{matrix} & 1 \end{array}\right]$$

For a block-triangular matrix $\begin{bmatrix} A & b \\ 0 & c \end{bmatrix}$, the inverse is still block-triangular, of the form $\begin{bmatrix} A^{-1} & d \\ 0 & c^{-1} \end{bmatrix}$

To find $d$, we impose that $\begin{bmatrix} A & b \\ 0 & c \end{bmatrix}\begin{bmatrix} A^{-1} & d \\ 0 & c^{-1} \end{bmatrix} = I$, which gives $d = -\frac{1}{c}A^{-1}b$

Therefore, in our case

$$Z_n^{-1} = \begin{bmatrix} Z_{n-1}^{-1} & -Z_{n-1}^{-1}\cdot\begin{bmatrix}-1\\-1\\\vdots\\-1\end{bmatrix} \\ 0 & 1 \end{bmatrix}$$

For $i<j\le n-1$ $\quad (Z_n^{-1})_{ij} = (Z_{n-1}^{-1})_{ij} = 2^{j-i-1}$ by inductive hypothesis

For $i<j=n$ we have

$$(Z_n^{-1})_{in} = \left(-Z_{n-1}^{-1}\cdot\begin{bmatrix}-1\\\vdots\\-1\end{bmatrix}\right)_i = \left(Z_{n-1}^{-1}\cdot\begin{bmatrix}1\\\vdots\\1\end{bmatrix}\right)_i =$$

$$= \sum_{k=i}^{n-1}(Z_{n-1}^{-1})_{ik} = 1 + \sum_{k=i+1}^{n-1} 2^{k-i-1} = 1 + \sum_{h=0}^{n-i-2} 2^h =$$

(use inductive hypothesis)

$$= 1 + \left(2^{n-i-1} - 1\right) = 2^{n-i-1} = 2^{j-i-1}. \qquad \square$$

## Problem 4

Wlog we assume $I = \{1, 2, \dots, r\}$. Then $\tilde{U} = \begin{bmatrix} I_r \\ B \end{bmatrix}$ for some $B \in \mathbb{R}^{(n-r)\times r}$.

As $\tilde{U} = U\cdot U(I,:)^{-1}$ we have that $\mathrm{Vol}(\tilde{U}(J,:)) = |\det(\tilde{U}(J,:))| =$

$$= |\det(U(J,:)\cdot U(I,:)^{-1})| = |\det(U(J,:))|\cdot|\det(U(I,:)^{-1})| =$$

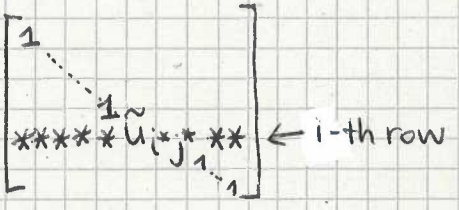
(Binet theorem)

$$= \mathrm{Vol}(U(J,:))\cdot\mathrm{Vol}(U(I,:)^{-1}) \quad \text{and similarly} \quad \mathrm{Vol}(\tilde{U}(I,:)) = \mathrm{Vol}(U(I,:))\cdot\mathrm{Vol}(U(I,:))$$

Therefore, it is sufficient to show that

$$\mathrm{Vol}(\tilde{U}(J,:)) = \tilde{u}_{i^*j^*}\cdot\mathrm{Vol}(\tilde{U}(I,:))$$

Note that $\tilde{U}(I,:) = \text{Identity}$ and $\tilde{U}(J,:) =$




therefore $|\det(\tilde{U}(J,:))| = |\tilde{u}_{i^*j^*}|$.

This proves that $\mathrm{Vol}(U(J,:)) = |\tilde{u}_{i^*j^*}|\cdot\mathrm{Vol}(U(I,:))$

In Knuth's algorithm, an exchange is made iff $|\tilde{u}_{i^*j^*}| \ge \mu$, so this means that the volume improves by a factor at least $\mu$. $\square$

## Problem 5

Recall that a projection $P$ is a matrix such that $P^2 = P$.
If $P$ is a projection, also $I-P$ is a projection: indeed, $(I-P)^2 = I - 2P + P^2 = I - 2P + P = I - P$. Therefore, it is sufficient to prove that for any projection $P$ it holds that $\|P\|_2 \leq \|I - P\|_2$.

Let $u \in \mathbb{R}^n$ be a unit vector such that $\|Pu\|_2 = \|P\|_2$.
We decompose $u = x + y$ where $x := Pu$ and $y := (I-P)u$.
We have $x \neq 0$ otherwise we would have $\|P\|_2 = 0$ which implies $P = 0$.
If $y = 0$, it means that $\|Pu\|_2 = 1 = \|P\|_2$; as $P \neq I$, $\operatorname{rank}(P) < n$ therefore $\exists$ (unit vector) $z \in \mathbb{R}^n$ s.t. $Pz = 0$, which gives $\|(I-P)z\|_2 = 1 \leq \|I - P\|_2$.

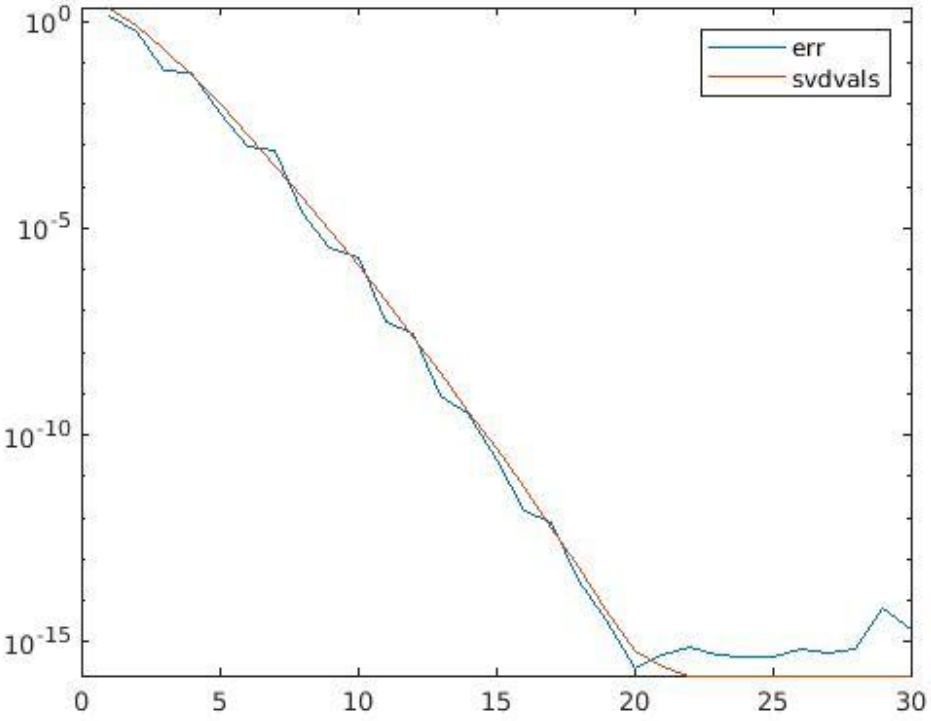
If $x \neq 0$ and $y \neq 0$, we define the vector $w := \frac{\|y\|_2}{\|x\|_2} x + \frac{\|x\|_2}{\|y\|_2} y$.
Note that $w$ is a unit vector, because

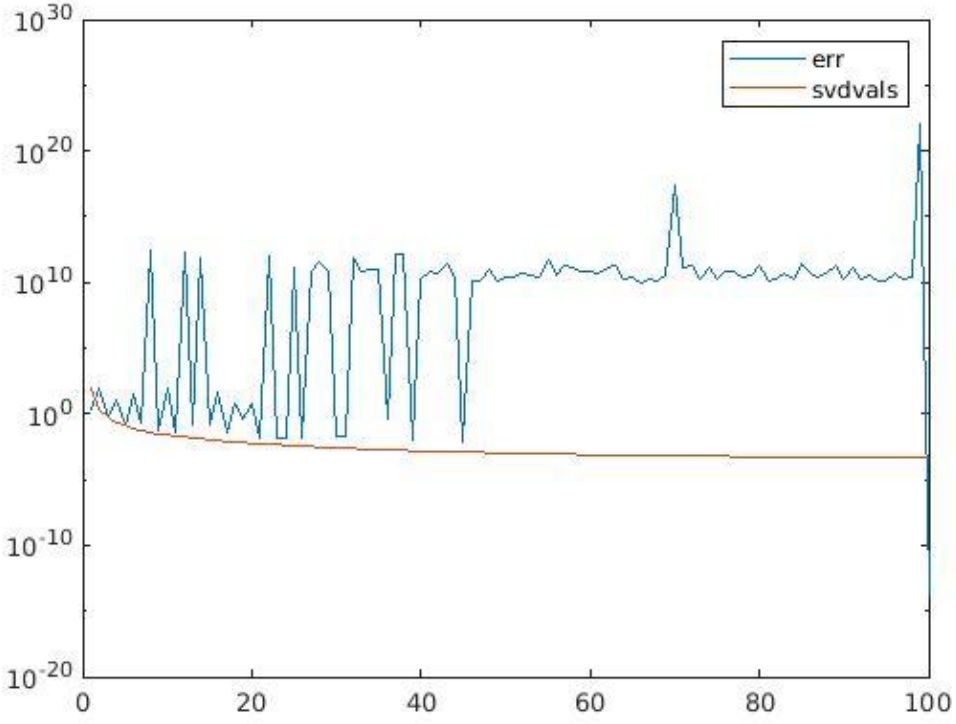
$$\|w\|^2 = \langle w, w \rangle = \frac{\|y\|_2^2}{\|x\|_2^2} \cdot \|x\|_2^2 + \frac{\|x\|_2^2}{\|y\|_2^2} \cdot \|y\|_2^2 + 2 \left\langle \frac{\|y\|_2}{\|x\|_2} x, \frac{\|x\|_2}{\|y\|_2} y \right\rangle =$$

$$= \|y\|_2^2 + \|x\|_2^2 + 2\langle x, y \rangle = \|u\|^2 = 1.$$

Moreover, we have that $Pw = \frac{\|y\|_2}{\|x\|_2} x$ and $(I-P)w = \frac{\|x\|_2}{\|y\|_2} y$, so $\|(I-P)w\|_2 = \|x\|_2$.
Therefore, $\|P\|_2 = \|Pu\|_2 = \|x\|_2 = \|(I-P)w\|_2 \leq \|I - P\|_2$. □

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

figure(1)
lowrank_approx(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)
lowrank_approx(A,R);


function lowrank_approx(A,R)
    [U,S,V] = svd(A,'econ');
    s = diag(S);
    err_greedy = zeros(R,1);
    for r=1:R
        Ur = U(:,1:r);
        I = greedy(Ur);
        Vr = V(:,1:r);
        J = greedy(Vr);
        Ar = A(:,J)/A(I,J)*A(I,:);
        err_greedy(r) = norm(A-Ar);
    end
    semilogy(1:R,err_greedy)
    hold on
    semilogy(1:R,s(1:R))
    legend("err","svdvals")
    hold off
end


function I = greedy(U)
    r = size(U,2);
    I = [];
    for k=1:r
        [~, imax] = max(abs(U(:,k)));
        imax = imax(1);
        I = [I; imax];
        U = U-1/U(imax,k)*U(:,k)*U(imax,:);
    end
end
```


err
svdvals
$10^{0}$
$10^{-5}$
$10^{-10}$
$10^{-15}$
0
5
10
15
20
25
30


err
svdvals
$10^{30}$
$10^{20}$
$10^{10}$
$10^{0}$
$10^{-10}$
$10^{-20}$
0
20
40
60
80
100