HINTS

Problem 1: write the product in terms of the entries of the matrices and use the linearity of $\mathbb{E}$. Remember that for $X,Y \sim N(0,1)$ independent, we have $\mathbb{E}[X^2] = 1$ and $\mathbb{E}[XY] = 0$.

Problem 2: 
1. Pay attention at the sizes of the matrices. How does your answer change if multiplication of $A \times $ vector takes less than $O(mn)$ time?
2. Apply triangular inequality to $\| A - QQ^T A QQ^T \|_2 = \| ( A - QQ^T A ) + ( QQ^T A - QQ^T A QQ^T ) \|_2$, use symmetry of $A$ and properties of the orthogonal projection $QQ^T$.
3. You want an approximation of the form $Q \cdot $ something $\cdot Q^T$.

Problem 3: 
For the first part, in addition to the given hints, recall that $\| PDP \|_2 = \max \frac{y^T PDP y}{y^T y}$: what happens for $y = P \cdot x$?
For the second part: consider $A = U \Sigma V^T$ and try to write $\| PA \|_2^{2(q+1)} = \| PAA^T P \|_2^{2q+1}$ in the form $\| \Sigma^2 \Sigma \|_2^{2q+1}$ for a suitable orthogonal projection $\Sigma$ and then apply the first part of the exercise.

Problem 4: 
To prove $\| P \|_2 Cw \|_2 \leq \| C \|_2$, write $C = U \Sigma V^T$ and $\| Cw \|_2 = \omega^T C^T C w$, recall that multiplying an orthogonal matrix with a Gaussian random vector $\sim N(0,1)$ gives a Gaussian random vector $\sim N(0,1)$. To prove $\| P \|_2 (|x| \leq \mu) \leq \sqrt{\frac{2}{\pi}} \mu$ write the LHS as the integral of the density and then estimate such integral in the easiest possible way.
PROBLEM 1

\[ E \left[ \| A \Omega B \|_F^2 \right] = E \left[ \sum_{i,j} (A \Omega B)_{i,j}^2 \right] = \sum_{i,j} E \left[ (A \Omega B)_{i,j}^2 \right] = \sum_{i,j} \left( \sum_{k,h} a_{i,k} \omega_{k,h} b_{h,j} \right)^2 \]

\[ = \sum_{i,j} \sum_{k,h} \left( a_{i,k}^2 \omega_{k,h}^2 b_{h,j}^2 \right) + \sum_{i,j} \left( \sum_{k,h} a_{i,k} a_{i,k} \omega_{k,h} \omega_{k,h} b_{h,j} b_{h,j} \right) \]

\[ = \sum_{i,j} \sum_{k,h} a_{i,k}^2 b_{h,j}^2 = \left( \sum_{i,j} a_{i,k}^2 \right) \left( \sum_{i,j} b_{h,j}^2 \right) = \| A \Omega \|_F^2 \cdot \| B \|_F^2. \]

\[ E[\omega_{k,h}^2] = 1, \ E[\omega_{k_1,h} \omega_{k_2,h}] = 0 \text{ for } (k_1,h_1) \neq (k_2,h_2) \]

Another proof: consider SVDs \( A = U_A \Sigma_A V_A^T \) and \( B = U_B \Sigma_B V_B^T \), then \( X := V_A^T \Sigma_A U_B \) is also Gaussian because \( V_A \) and \( U_B \) are orthogonal matrices. We have \( \| A \Omega B \|_F = \| U_A \Sigma_A X \Sigma_B V_B^T \|_F = \| \Sigma_A X \Sigma_B \|_F \)

therefore

\[ E \left[ \| A \Omega B \|_F^2 \right] = E \left[ \| \Sigma_A X \Sigma_B \|_F^2 \right] = E \left[ \sum_{i,j} (\sigma_A^2)^2 (\sigma_B^2)^2 \right] = \left( \sum_{i,j} (\sigma_A^2)^2 \right) \left( \sum_{i,j} (\sigma_B^2)^2 \right) = \| A \Omega \|_F^2 \cdot \| B \|_F^2. \]

PROBLEM 2

\( \text{①} \) The multiplication \( Y = A \Omega \) costs \( O(mn(r+p)) \), the economy QR factorization costs \( O(m(r+p)^2) \), computation of \( B \) costs \( O(nm(r+p)) \) and computing \( T_r(B) \) costs \( O(n(r+p)^2) \) so the total cost is \( O(mn(r+p)) \). Note that if we call \( C_{\text{mult}} \) the cost of one matrix-vector multiplication with \( A \) (which can be lower than \( O(mn) \) if \( A \) has some special structure), then the cost of the algorithm is \( O(C_{\text{mult}} \cdot (r+p) + (m+n)(r+p)^2) \).

\( \text{②} \) \( \| A - QQ^T A Q Q^T \|_2 = \| A - QQ^T A + QQ^T A - QQ^T A Q Q^T \|_2 \)

\[ \leq \| A - QQ^T A \|_2 + \| QQ^T (A - A Q Q^T) \|_2 \leq E + \| QQ^T \|_2 \| A - A Q Q^T \|_2 \]

triangular inequality \( = E + \| QQ^T \|_2 \cdot \| A - A Q Q^T \|_2 \leq 2E. \)

\( \text{③} \) Instead of \( B = Q^T A \), compute \( B = Q^T A Q \) (which is symmetric!), find a spectral decomposition of this small matrix: \( B \approx U \Lambda U^T \), and take \( A \approx Q \tilde{U} \cdot \Lambda \cdot (Q \tilde{U})^T = U \Lambda U^T \).
PROBLEM 3

First we prove that \( \| P D P \|_2 \leq \| P D^e P \|_2 \). Let \( x = (x_1, \ldots, x_n)^T \) be a unit vector such that \( x^T P D P x = \| P D P \|_2 \). Assume by contradiction that \( P x \neq x \); as \( P \) is an orthogonal projection, \( \| P x \|_2 < 1 \). Then the vector \( y := P x \) is such that

\[
\frac{y^T P D P y}{y^T y} = \frac{x^T P^2 D^2 x}{\| P x \|_2^2} = \frac{\| P D P \|_2^2}{\| P x \|_2^2} > \| P D P \|_2^2.
\]

which is a contradiction as \( \| P D P \|_2^2 = \max_{x \neq 0} \frac{x^T P D P x}{x^T x} \).

Therefore \( P x = x \). So

\[
\| P D P \|_2^t = (x^T (P D P) x)^t = (x^T D x)^t = \left( \sum_{i=1}^n d_i x_i^2 \right)^t \leq \sum_{i=1}^n d_i x_i^2 = x^T D x = (P x)^T D (P x) \leq \| P D P \|_2^t.
\]

Jensen's inequality for convex function \( z \mapsto |z|^t \) (using that \( d_i \geq 0 \) \( \forall i \) and \( \sum x_i^2 = 1 \))

We now prove the second part of the exercise.

We have that \( \| PA \|_2^{2q+1} = \| P A A^T P \|_2^{2q+1} = \| P U Z^2 U^T P \|_2^{2q+1} = \| (U^T P U) \Sigma^2 (U^T P U) \|_2^{2q+1} = \| (U^T P U) \Sigma^{2(2q+1)} (U^T P U) \|_2^{2q+1} \) = \( \| P(A A^T)^{2q} P \|_2 = \| P(A A^T)^q A A^T(A A^T)^q P \|_2 \) = \( \| A \|_2^2 \).

PROBLEM 4

Let \( C = U \Sigma V^T \) be an SVD of \( C \), then

\[
P \left( \frac{\| C w \|_2 \leq \mu \| C \|_2}{} = \frac{\| w^T C^T C w \leq \mu^2 \sigma_1^2}{} = \frac{\| w^T V \Sigma^2 V^T w \leq \mu^2 \sigma_1^2}{} = \frac{\| X^T \Sigma^2 X \leq \mu^2 \sigma_1^2}{} \right)
\]

\( X \sim N(0, I) \) is also a Gaussian vector.

\[
P \left( \sum_{i=1}^n \sigma_i^2 x_i^2 \leq \mu^2 \sigma_1^2 \right) \leq \frac{4}{\sqrt{2\pi}} \cdot 2\mu = \frac{4\sqrt{2}}{\sqrt{2\pi}} \mu.
\]

Choosing \( \mu = \frac{1}{10} \sqrt{\frac{\pi}{2}} \) gives

\[
P \left( \| A \omega \|_2 \leq \frac{4}{10} \sqrt{\frac{\pi}{2}} \| A \|_2 \right) \leq \frac{4}{10}.
\]

Therefore,

\[
P \left( \max_{i=1, \ldots, \kappa} \| A \omega_i \|_2 \leq \frac{4}{10} \sqrt{\frac{\pi}{2}} \| A \|_2 \right) = \frac{\kappa}{\kappa} P \left( \| A \omega_i \|_2 \leq \frac{4}{10} \sqrt{\frac{\pi}{2}} \| A \|_2 \right) \leq \left( \frac{4}{10} \right)^\kappa.
\]

This means that \( P \left( \| A \|_2 \leq 10 \sqrt{\frac{\pi}{2}} \max_{i=1, \ldots, \kappa} \| A \omega_i \|_2 \right) \geq 1 - \frac{4}{10} \kappa \).
clear all
n=100;
x=zeros(n);
for i=1:n
    x(i)=i/10;
end
A=zeros(n,n);
for i=1:n
    for j=1:n
        A(i,j)=1./(x(i)+x(j));
    end
end
B=zeros(n,n);
for i=1:n
    for j=1:n
        B(i,j)=1./((x(i)+x(j))^0.5);
    end
end
rA=rank(A)
rB=rank(B)
[Ua,Sa,Va]=svd(A, 'econ');
[Ub,Sb,Vb]=svd(B, 'econ');

Omega=randn(n,k);
Ya=Ua*(repmat(diag(Sa),1,k).*Va'*Omega));
Yb=Ub*(repmat(diag(Sb),1,k).*Vb'*Omega));
Yc=Ya+Yb;

[Q,R]=qr(Yc,0);
Ba=Q'*Ua*(repmat(diag(Sa),1,n).*Va');
Bb=Q'*Ub*(repmat(diag(Sb),1,n).*Vb');
B=Ba+Bb;
[U,S,V]=svd(B, 'econ');

Published with MATLAB® R2020a