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 $\|A - QQ^TAQQ^T\|_2 = \|(A - QQ^TA) + (QQ^TA - QQ^TAQQ^T)\|_2$, use symmetry of A and properties of the orthogonal projection QQ^T.

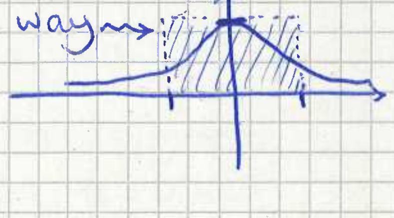
③ You want an approximation of the form $Q \cdot$ something $\cdot Q^T$.

Problem 3: For the first part, in addition to the given hints, recall that $\|PDP\|_2 = \max_{y \neq 0} \frac{y^T PDP y}{y^T y}$: what happens for $y := P \cdot x$?

For the second part: consider $A = U \Sigma V^T$ and try to write $\|PA\|_2^{2(2q+1)} = \|PAA^TP\|_2^{2q+1}$ in the form " $\|\tilde{P} \Sigma^2 \tilde{P}\|_2^{2q+1}$ for a suitable orthogonal projection $\tilde{P}$ and then apply the first part of the exercise.

Problem 4: To prove $\mathbb{P}(\|Cw\|_2 \leq \mu \|C\|_2)$ write $C = U \Sigma V^T$ and $\|Cw\|_2^2 = w^T C^T C w$, recall that multiplying an orthogonal matrix with a Gaussian random vector $\sim N(0, I)$ gives a Gaussian random vector $\sim N(0, I)$.

To prove $\mathbb{P}(|X| \leq \mu) \leq \sqrt{\frac{2}{\pi}} \mu$ write the LHS as the integral of the density and then estimate such integral in the easiest possible way ⇝



SOLUTIONS

PROBLEM 1

$$\mathbb{E}\left[\|A\Omega B\|_F^2\right] = \mathbb{E}\left[\sum_{i,j}(A\Omega B)_{ij}^2\right] = \sum_{i,j}\mathbb{E}\left[(A\Omega B)_{ij}^2\right] =$$

$$= \sum_{i,j}\mathbb{E}\left[\left(\sum_{k,h} a_{ik}\,\omega_{kh}\, b_{hj}\right)^2\right] =$$

$$= \sum_{i,j}\left(\sum_{k,h}\mathbb{E}\left[a_{ik}^2\omega_{kh}^2 b_{hj}^2\right] + \sum_{(k_1,h_1)\neq(k_2,h_2)}\mathbb{E}\left[a_{ik_1}a_{ik_2}\omega_{k_1h_1}\omega_{k_2h_2}b_{h_1j}b_{h_2j}\right]\right)$$

$$\underset{\uparrow}{=} \sum_{i,j}\sum_{k,h} a_{ik}^2 b_{hj}^2 = \left(\sum_{i,k}a_{ik}^2\right)\cdot\left(\sum_{h,j}b_{hj}^2\right) = \|A\|_F^2\cdot\|B\|_F^2.$$

$\mathbb{E}[\omega_{kh}^2]=1$, $\mathbb{E}[\omega_{k_1h_1}\omega_{k_2h_2}]=0$ for $(k_1,h_1)\neq(k_2,h_2)$ $\square$

Another proof: consider SVDs $A=U_A\Sigma_A V_A^T$ and $B=U_B\Sigma_B V_B^T$, then $X := V_A^T\Omega U_B$ 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

$$\mathbb{E}\left[\|A\Omega B\|_F^2\right] = \mathbb{E}\left[\|\Sigma_A X\Sigma_B\|_F^2\right] = \mathbb{E}\left[\sum_{i,j}(\sigma_i^A)^2(\sigma_j^B)^2 X_{ij}^2\right] \overset{\mathbb{E}[X_{ij}^2]=1}{=}$$

$$= \sum_{i,j}(\sigma_i^A)^2(\sigma_j^B)^2 = \left(\sum_i(\sigma_i^A)^2\right)\cdot\left(\sum_j(\sigma_j^B)^2\right) = \|A\|_F^2\cdot\|B\|_F^2. \qquad \square$$

PROBLEM 2

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_{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 $O\left(C_{mult}\cdot(r+p) + (m+n)(r+p)^2\right)$.

② $$\|A - QQ^TAQQ^T\|_2 = \|A - QQ^TA + QQ^TA - QQ^TAQQ^T\|_2$$

$$\underset{\substack{\uparrow\\ \text{triangular inequality}}}{\leq} \|A-QQ^TA\|_2 + \|QQ^T(A-AQQ^T)\|_2 \leq \varepsilon + \|QQ^T\|_2\,\|A-AQQ^T\|_2$$

$$= \varepsilon + \|QQ^T\|_2\cdot\|A^T - QQ^TA^T\|_2 =$$

$$= \varepsilon + \underbrace{\|QQ^T\|_2}_{\leq 1}\cdot\|A-QQ^TA\|_2 \leq 2\varepsilon. \qquad \square$$

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

[SOLUTIONS]

PROBLEM 3

First we prove that $\|PDP\|_2^t \le \|PD^tP\|_2$. Let $x=(x_1,\dots,x_n)^T$ be a unit vector such that $x^TPDPx=\|PDP\|_2$. Assume by contradiction that $Px\neq x$; as P is an orthogonal projection, $\|Px\|_2<1$. Then the vector $y:=Px$ is such that

$$\frac{y^TPDPy}{y^Ty}=\frac{x^TP^2DP^2x}{\|Px\|_2^2}=\frac{x^TPDPx}{\|Px\|_2^2}=\frac{\|PDP\|_2}{\|Px\|_2^2}>\|PDP\|_2$$

which is a contradiction as $\|PDP\|_2=\max_{z\neq 0}\frac{z^TPDPz}{z^Tz}$.

Therefore $Px=x$. So

$$\|PDP\|_2^t=\left(x^T(PDP)x\right)^t=\left(x^TDx\right)^t=\left(\sum_{i=1}^n d_i x_i^2\right)^t \le$$

$$\le \sum_{i=1}^n d_i^t x_j^2 = x^TD^tx=(Px)^TD^t(Px)$$

$$\le \|PD^tP\|_2.$$

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

We now prove the second part of the exercise.

We have that $\|PA\|_2^{(2q+1)\cdot 2}=\|PAA^TP\|_2^{2q+1}\underset{\substack{\uparrow\\ \text{let } A=U\Sigma V^T \text{ be SVD}}}{=}\|PU\Sigma^2U^TP\|_2^{2q+1}=$

$$=\|\underbrace{(U^TPU)}_{\text{orthogonal projection}}\overset{\text{diag. matrix with nonneg. entries}}{\Sigma^2}(U^TPU)\|_2^{2q+1}\le\|(U^TPU)\Sigma^{2(2q+1)}(U^TPU)\|_2=$$

$$=\|P(AA^T)^{2q+1}P\|_2=\|P(AA^T)^qAA^T(AA^T)^qP\|_2=$$

$$=\|P(AA^T)^qA\|_2^2. \qquad \square$$

PROBLEM 4

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

$$\mathbb{P}\left(\|C\omega\|_2\le\mu\|C\|_2\right)=\mathbb{P}\left(\omega^TC^TC\omega\le\mu^2\sigma_1^2\right)=$$

$$=\mathbb{P}\left(\omega^TV\Sigma^2V^T\omega\le\mu^2\sigma_1^2\right)=\mathbb{P}\left(X^T\Sigma^2X\le\mu^2\sigma_1^2\right)=$$

denote $X=(X_1,\dots,X_n)$, $X_i\sim N(0,1)\ \forall i$

$X:=V^T\omega\sim N(0,I)$ is also a Gaussian vector

$$=\mathbb{P}\left(\sum_{i=1}^n\sigma_i^2X_i^2\le\mu^2\sigma_1^2\right)\le\mathbb{P}\left(\cancel{\sigma_1^2}X_1^2\le\cancel{\sigma_1^2}\mu^2\right)=\frac{1}{\sqrt{2\pi}}\int_{-\mu}^{\mu}\exp\left(-\tfrac12 t^2\right)dt$$

$\le\frac{1}{\sqrt{2\pi}}\cdot 2\mu=\sqrt{\frac{2}{\pi}}\mu$. Choosing $\mu:=\frac{1}{10}\sqrt{\frac{\pi}{2}}$ gives

$\mathbb{P}\left(\|A\omega\|_2\le\frac{1}{10}\sqrt{\frac{\pi}{2}}\|A\|_2\right)\le\frac{1}{10}$. Therefore,

$$\mathbb{P}\left(\max_{i=1,\dots,k}\|A\omega^{(i)}\|_2\le\frac{1}{10}\sqrt{\frac{\pi}{2}}\|A\|_2\right)=\prod_{i=1}^k\mathbb{P}\left(\|A\omega^{(i)}\|_2\le\frac{1}{10}\sqrt{\frac{\pi}{2}}\|A\|_2\right)\le\left(\frac{1}{10}\right)^k.$$

This means that $\mathbb{P}\left(\|A\|_2\le 10\sqrt{\frac{2}{\pi}}\max_{i=1,\dots,k}\|A\omega^{(i)}\|_2\right)\ge 1-\frac{1}{10^k}$. $\square$

```
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');
Ua=Ua(:,1:rA);
Sa=Sa(1:rA,1:rA);
Va=Va(:,1:rA);
Ub=Ub(:,1:rB);
Sb=Sb(1:rB,1:rB);
Vb=Vb(:,1:rB);
C=A+B;
r=rank(C)
k=r+0;
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');
U=Q*U(:,1:r);
S=S(1:r,1:r);
V=V(:,1:r);
norm(C-U*S*V')
sv=svds(C,r+1);
sv(end)
```

Published with MATLAB® R2020a