MATH-605: Homework # 4

Due on Tuesday, October 31, 2017

Saket Choudhary
2170058637

Contents

5.4.12

$$
\begin{aligned}
\mathbb{E}\exp\lambda\epsilon A &= \frac{1}{2}(\exp\lambda A + \exp -\lambda A)\\
\exp A &= I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots\\
\exp -A &= I - A + \frac{A^2}{2!} - \frac{A^3}{3!} + \dots\\
\frac{1}{2}(\exp\lambda A + \exp -\lambda A) &= I + \frac{\lambda^2 A^2}{2} + \frac{\lambda^4 A^4}{4!} + \frac{\lambda^6 A^6}{6!} + \dots\\
&= I + \frac{\lambda^2 A^2}{2} + \frac{\lambda^4 A^4}{4*3} + \frac{\lambda^6 A^6}{8*90} + \dots\\
&\le 1 + (\lambda^2 A^2/2) + \frac{((\lambda^2 A^2)/2)^2}{2} + \frac{(\lambda^2 A^2/2)^3}{3!} + \dots\\
&= \exp\lambda^2 A^2/2
\end{aligned}
$$

Then define $X = \sum_{i=1}^N \epsilon_i A_i$, then following from 5.14 :

$$
\begin{aligned}
\mathbb{P}\{\lambda_{max}(S) \ge t\} &= \mathbb{P}\{e^{\lambda\lambda_{max}(S)\ge t}\}\\
&\le e^{-\lambda t}\mathbb{E}e^{\lambda.\lambda_{max}(S)}
\end{aligned}
$$

Define $E = \mathbb{E}\lambda_{max}(e^{\lambda S})$. By the bound on maximum eigen value of $e^{\lambda S}$: $E \le \mathbb{E}tre^{\lambda S}$
Applying Lieb's inequality:

$$
\begin{aligned}
&E \le \mathbb{E}\ tre^{\lambda S}\\
&E \le \mathbb{E}\ tr\exp[\sum_{i=1}^{N-1}\lambda X_i + \lambda X_N]\\
&E \le \mathbb{E}\ tr\exp[\sum_{i=1}^{N-1}\lambda X_i + \log\mathbb{E}e^{\lambda X_N}] && \text{Conditioning on } X_{i=1}^{N-1} \text{ and using Lemma 5.4.9}\\
&E \le tr\exp[\sum_{i=1}^{N}\log\mathbb{E}e^{\lambda X_i}] && \text{Repeating above step and using Lemma 5.4.9 } N \text{ times}\\
&\quad \le tr\exp[\sum_{i=1}^{N}\frac{\lambda_i^2 A_i^2}{2}] && \because \mathbb{E}\exp\lambda\epsilon A \le \exp\lambda^2 A^2/2\\
&\quad \le n.\lambda_{max}(\exp\sum_{i=1}^{N}\lambda_i^2 A_i^2/2) && \because \text{trace is sum of } n \text{ eigen values}\\
&\quad \le n\exp\{n\lambda_{max}^2/2\sum_{i=1}^{n}A_i^2\} && \because \lambda_i \le \lambda_{max}\forall i \in [1,N]\\
&\quad = n\exp\{\lambda_{max}^2\sigma^2/2\}\\
&\mathbb{P}\{\lambda_{max}(S) \ge t\} \le n.\exp\{-\lambda t + \lambda_{max}^2\sigma^2/2\} && \text{substituting for E}
\end{aligned}
$$

Differentiating $\exp\{-\lambda t + \lambda_{max}^2\sigma^2/2\}$ wrt λ gives: $\exp\{-\lambda t + \lambda_{max}^2\sigma^2/2\} * (-t + \lambda_{max}\sigma^2/2)$ Thus $\lambda = \frac{t}{\sigma^2}$ and hence:

$$
E \le n.\exp\{-\frac{t^2}{2\sigma^2}\}
$$

I don't seem to get a factor of 2 on the RHS

5.4.15

Consider dilation of X as $Y := \begin{pmatrix} 0 & X_i^T \\ X_i & 0 \end{pmatrix}$ Then $Y^2 = \begin{pmatrix} X_i^T X_i & 0 \\ 0 & X_i X_i^T \end{pmatrix}$ Then $\sigma^2 = ||\sum_{i=1}^N Y_i^2|| = ||\begin{pmatrix} \sum_{i=1}^N X_i^T X_i & 0 \\ 0 & \sum_{i=1}^N X_i X_i^T \end{pmatrix}|| = \max\{||\sum_{i=1}^N X_i^T X_i||, ||\sum_{i=1}^N X_i X_i^T||\}$ Applying Matrix Bernstein's inequality from theorem 5.4.1 to the dilation Y of X give:

$$
\begin{aligned}
P\{||\sum_{i=1}^N X_i|| \geq t\} &= P\{||\sum_{i=1}^N Y_i|| \geq t\} \\
&\leq 2(m+n)\exp - \frac{t^2/2}{\sigma^2 + Kt/3}
\end{aligned}
$$

where $\sigma^2 = \max\{||\sum_{i=1}^N X_i^T X_i||, ||\sum_{i=1}^N X_i X_i^T||\}$

5.6.6

Frame u_i obeys's approximate Parseva;s identity: $\exists A, B > 0$ such that

$$
A||x||_2^2 \leq \sum_{i=1}^N \langle u_i, x\rangle \leq B||x||_2^2 \forall x \in \mathbb{R}^n
$$

u_i is tight when $A = B$. Also, from problem 3.3.9 we have $\{u_i\}_{i=1}^N$ is tight when $\sum_{i=1}^n u_i u_i^T = AI_n$
Consider random sample $\{v_i\}_{i=1}^m$ of $\{u_i\}$ from remark 5.6.2 we see that

$$
\begin{aligned}
E||\sum_{i=1}^m v_i v_i^T - \sum_{i=1}^N u_i u_i^T|| &\leq \epsilon||\sum_{i=1}^N u_i u_i^T|| \\
&= \epsilon||A||
\end{aligned}
$$

Hence $\{v_i\}_{i=1}^m$ has a good frame bound.

6.1.6

$EF(\sum_{i\neq j} a_{ij} f(X_i, X_j)) \leq E(4\sum_{i\neq j} a_{ij} f(X_i, X_j'))$. For this to hold, f should be measurable.
If this holds then theorem 6.1.1 is implied by taking $f(X_i, X_j) = X_i X_j$ for matrix X_i and theorem 6.1.4 is implied by considering $f(X_i, X_j) = X_i X_j^T$ for vectors X_i, X_j.

6.3.4

To prove: $\mathbb{E}||\sum_{i=1}^{N} X_i - \sum_{i=1}^{N} \mathbb{E}X_i|| \leq 2\mathbb{E}||\sum_{i=1}^{N} \epsilon_i X_i||$

$$
\begin{aligned}
\mathbb{E}||\sum_{i=1}^{N} X_i - \sum_{i=1}^{N} EX_i|| &\leq \mathbb{E}||\sum_{i=1}^{N} X_i|| + ||\sum_{i=1}^{N} \mathbb{E}X_i|| && \text{triangular inequality} \\
&\leq \mathbb{E}||\sum_{i=1}^{N} \epsilon_i X_i|| + ||\sum_{i=1}^{N} \mathbb{E}\epsilon_i X_i|| && \text{same distribution} \\
&\leq \mathbb{E}||\sum_{i=1}^{N} \epsilon_i X_i|| + \mathbb{E}||\sum_{i=1}^{N} \epsilon_i X_i|| && \text{Jensen's inequality} \\
&= 2\mathbb{E}||\sum_{i=1}^{N} \epsilon_i X_i||
\end{aligned}
$$

6.3.5

Consider Y_i to be independent copies of X_i

$$
\begin{aligned}
||\sum_{i=1}^{N} X_i||_F &= ||\sum_{i=1}^{N} X_i - \sum_{i=1}^{N} \mathbb{E}Y_i||_F && \because EY_i = 0 \\
&\leq \mathbb{E}_Y||\sum_{i=1}^{N}(X_i - Y_i)||_F \\
F(||\sum_{i=1}^{N} X_i||_F) &\leq F(\mathbb{E}_Y||\sum_{i=1}^{N}(X_i - Y_i)||_F) \\
&\leq \mathbb{E}_Y F(||\sum_{i=1}^{N}(X_i - Y_i)||_F) && \text{Jensen's} \\
\mathbb{E}_{X,Y} F(||\sum_{i=1}^{N} X_i||_F) &\leq \mathbb{E}_X\mathbb{E}_Y F(||\sum_{i=1}^{N}(X_i - Y_i)||_F) \\
&\leq \mathbb{E}_{X,Y} F(||\sum_{i=1}^{N}(X_i - Y_i)||_F) && \text{Using Fubini's}
\end{aligned}
$$

Now, for rademacher ϵ_i

$$
\begin{aligned}
\mathbb{E}_{X,Y} F(||\sum_{i=1}^{N}(X_i - Y_i)||_F &\leq \mathbb{E}_\epsilon\mathbb{E}_{X,Y} F(||\sum_{i=1}^{N}\epsilon_i(X_i - Y_i)||_F) \\
&\leq \mathbb{E}_{X,Y} F(||\sum_{i=1}^{N}\epsilon_i X_i||_F) + \mathbb{E}_{X,Y} F(||\sum_{i=1}^{N}\epsilon_i Y_i||_F \\
&\leq 2\mathbb{E}_{X,Y} F(||\sum_{i=1}^{N}\epsilon_i X_i||_F)
\end{aligned}
$$

which is the upper bound.
For lower bound:

$$
\begin{aligned}
\mathbb{E}F(\frac{1}{2}||\sum_{i=1}^{N}\epsilon_i X_i||_F) &= \mathbb{E}_\epsilon\mathbb{E}F(\frac{1}{2}||\sum_{i=1}^{N}\epsilon_i X_i - \sum_{i=1}^{N}\mathbb{E}Y_i||_F) \\
&\leq \mathbb{E}||\frac{1}{2}\sum_{i=1}^{N}\epsilon_i(X_i - Y_i)||_F \\
&\leq \mathbb{E}||\frac{1}{2}\sum_{i=1}^{N}\epsilon_i X_i||_F + \mathbb{E}||\frac{1}{2}\sum_{i=1}^{N}\epsilon_i X_i||_F \\
&= \mathbb{E}||\frac{1}{2}\sum_{i=1}^{N} X_i||_F + \mathbb{E}||\frac{1}{2}\sum_{i=1}^{N} Y_i||_F && \text{Same distrubution} \\
&\leq \frac{1}{2}(\mathbb{E}||\sum_{i=1}^{N} X_i||_F + \mathbb{E}||\sum_{i=1}^{N} Y_i||_F) \\
&= \mathbb{E}||\sum_{i=1}^{N} X_i||_F
\end{aligned}
$$

which is the LHS of the whole inequality.

6.3.5 continued on next page. . .

6.5.4

$\hat{X}$ is best approximation to $p^{-1}Y$ hence $||\hat{X} - p^{-1}Y|| \le ||p^{-1}Y - X||$

$$\begin{aligned} ||\hat{X} - X|| &\le ||\hat{X} - p^{-1}Y|| + ||p^{-1}Y - X|| && \text{Trinagular inequality} \\ ||\hat{X} - X|| &\le 2||p^{-1}Y - X|| && \because \text{Assumption above} \\ &= \frac{2}{p}||Y - pX|| \\ (Y - pX)_{ij} &= (\delta_{ij} - p)X_{ij} + \delta_{ij}v_{ij} \end{aligned}$$

Consider,

$$\begin{aligned} ||(Y - pX)_i||_2^2 &= \sum_{j=1}^{n}(\delta_{ij} - p)X_ij + \delta_{ij}v_{ij})^2 \\ &\le \sum_{j=1}^{n}((\delta_{ij} - p)||X||_\infty + \delta_{ij}||v||_\infty)^2 \\ \mathbb{E}\max\sum_{j=1}(\delta_{ij} - p)^2 &\le Cpn\text{Using 2.8.3} \\ \mathbb{E}\max\sum_{j=1}(\delta_{ij})^2 &\le Cpn \\ \implies \mathbb{E}\max||(Y - pX)_i||_2^2 &\le \mathbb{E}\sum_{j=1}^{n}((\delta_{ij} - p)||X||_\infty + \delta_{ij}||v||_\infty)^2 \\ &= \mathbb{E}\sum_{j=1}^{n}((\delta_{ij} - p)^2||X||_\infty^2 + \delta_{ij}^2||v||_\infty^2 + (\delta_{ij} - p)\delta_{ij}||X||_\infty||v||_\infty)^2 \\ \implies \mathbb{E}\max||(Y - pX)_i||_2 &\le \sqrt{pn}||X||_\infty + \sqrt{pn}||v|_\infty \end{aligned}$$

Using 6.4.2 .

$$\mathbb{E}||(Y - pX)|| \le C\sqrt{\log n}(\mathbb{E}\max||(Y - pX)_i|| + \mathbb{E}\max||(Y - pX)^j||)$$

Thus,

$$\mathbb{E}||(\hat{X} - X)|| \le \sqrt{\frac{n\log n}{p}}||X||_\infty + \sqrt{\frac{n\log n}{p}}||v||_\infty$$

6.6.5

$$
\begin{aligned}
\mathbb{E}||g||_\infty &= \mathbb{E}\{\max_{i\leq n} g_i\} \\
e^{s\mathbb{E}[\max_{i\leq n} g_i]} &\leq \mathbb{E}[e^{s\max_{i\leq n} g_i}] && \text{Jensen's inequality} \\
&= \mathbb{E}[\max e^{sg_i}] \\
&\leq \sum_{i=1}^{N} \mathbb{E}[e^{sg_i}] \\
&\leq ne^{\sigma^2 s^2/2} && \text{Using mgf of } g \\
\implies \mathbb{E}\{\max_{i\leq n} g_i\} &\leq \frac{\ln}{s} + \frac{s\sigma^2}{2}
\end{aligned}
$$

Differentiating $\frac{\ln}{s} + \frac{s\sigma^2}{2}$ wrt s gives $s = \sqrt{\frac{2\ln n}{\sigma^2}}$ and hence

$$
\begin{aligned}
\mathbb{E}||g||_\infty &= \mathbb{E}\{\max_{i\leq n} g_i\} \\
&\leq \sqrt{2}\sigma\sqrt{\ln n}
\end{aligned}
$$