# Homework 3

Benjamin Noland

1. Since $a \leq X \leq b$, we have the inequality

$$a + \frac{-b-a}{2} \leq X + \frac{-b-a}{2} \leq b + \frac{-b-a}{2},$$

which upon simplification becomes

$$\frac{a-b}{2} \leq X + \frac{-b-a}{2} \leq \frac{b-a}{2}.$$

Thus,

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{Var}\left(X + \frac{-b-a}{2}\right) \\
&= \mathrm{E}\left[\left(X + \frac{-b-a}{2}\right)^2\right] - \mathrm{E}\left[\left(X + \frac{-b-a}{2}\right)\right]^2 \\
&\leq \mathrm{E}\left[\left(X + \frac{-b-a}{2}\right)^2\right] \leq \left(\frac{b-a}{2}\right)^2 = \frac{(b-a)^2}{4}.
\end{aligned}$$

However, this inequality is not tight. To see this, let $X \sim \mathrm{unif}(0,1)$. Then $0 \leq X \leq 1$, but

$$\mathrm{Var}(X) = \frac{1}{12} < \frac{1}{4}.$$

2. (i) First, suppose $p \geq 1$ is even. Then the statement to prove reduces to

$$\Gamma(k) \leq 3k^k \quad \text{for every integer } k \geq 1. \tag{1}$$

I will proceed by induction on $k$. Since $\Gamma(1) = 1 \leq 3$, we see that (1) holds when $k = 1$. Now suppose that (1) holds for some $k > 1$. Then

$$\Gamma(k+1) = k\Gamma(k) \leq k \cdot 3k^k = 3k^{k+1} \leq 3(k+1)^{k+1}.$$

Thus (1) holds for every $k \geq 1$ by induction.

Next, suppose that $p \geq 1$ is odd. Then the statement to prove reduces to

$$\Gamma(k + 1/2) \leq 3(k+1/2)^{k+1/2} \quad \text{for every integer } k \geq 0. \tag{2}$$

As before, I will proceed by induction on $k$. Since $\Gamma(1/2) = \sqrt{\pi} \le 3(1/2)^{1/2}$, we see that (2) holds when $k = 0$. Now suppose that (2) holds for some $k > 0$. Then

$$\begin{aligned}
\Gamma((k+1)+1/2) &= (k+1/2)\Gamma(k+1/2) \\
&\le (k+1/2)\cdot 3(k+1/2)^{k+1/2} \\
&= 3(k+1/2)^{(k+1)+1/2} \\
&\le 3((k+1)+1/2)^{(k+1)+1/2}.
\end{aligned}$$

Thus (2) holds for every $k \ge 0$ by induction.

Combining these two cases, we therefore see that

$$\Gamma(p/2) \le 3(p/2)^{p/2} \quad \text{for every integer } p \ge 1.$$

(ii) Define a map $f : [0,\infty) \to \mathbb{R}$ by $f(x) = (2\sqrt{2})^x - x$. Note that $f(1) = 2\sqrt{2} - 1 > 0$. Moreover, since $\ln(2\sqrt{2}) > 1$, $f'(x) = \ln(2\sqrt{2})(2\sqrt{2})^x - 1 > 0$ for every $x \ge 1$. Thus $f$ is strictly increasing on $[1,\infty)$. In particular, since $f(1) > 0$ as noted, $f(x) > 0$ for every $x \ge 1$. Equivalently, $(2\sqrt{2})^x > x$ for every $x \ge 1$. Thus, for every integer $p \ge 1$, $(2\sqrt{2})^p > p$, so that $p^{1/p} \le 2\sqrt{2}$.

3. (b) $\implies$ (c). Let $p \ge 1$. Then

$$\begin{aligned}
\mathrm{E}(|X|^p) &= \int_0^\infty \mathrm{P}\{|X|^p > u\}\,du \\
&= \int_0^\infty \mathrm{P}\{|X|^p > t^p\} p t^{p-1}\,dt \quad (\text{substitute } u = t^p) \\
&= \int_0^\infty \mathrm{P}\{|X| > t\} p t^{p-1}\,dt \\
&\le \int_0^\infty 2e^{-t^2/K_1} p t^{p-1}\,dt \quad (\text{by (b)}) \\
&= \int_0^\infty e^{-s} p (K_1 s)^{p/2-1} K_1\,ds \quad (\text{subsitute } s = t^2/K_1) \\
&= pK_1^{p/2} \int_0^\infty e^{-s} s^{p/2-1}\,ds \\
&= pK_1^{p/2}\Gamma(p/2) \\
&\le 3K_1^{p/2} p (p/2)^{p/2} \quad (\text{by problem 2(i)}).
\end{aligned}$$

Thus,

$$\begin{aligned}
[\mathrm{E}(|X|^p)]^{1/p} &\le 3^{1/p} K_1^{1/2} p^{1/p} (p/2)^{1/2} \\
&\le 3K_1^{1/2} p^{1/p} (p/2)^{1/2} \quad (\text{since } 3^{1/p} \le 3) \\
&\le 6\sqrt{2} K_1^{1/2} (p/2)^{1/2} \quad (\text{by problem 2(ii)}) \\
&= 6K_1^{1/2}\sqrt{p}.
\end{aligned}$$

Thus, letting $K_2 = 6K_1^{1/2} > 0$, we have

$$[\mathrm{E}(|X|^p)]^{1/p} \leq K_2\sqrt{p}.$$

(c) $\implies$ (d). Let $K_3 = 4eK_2^2 > 0$. We have the following:

$$\begin{aligned}
\mathrm{E}(e^{X^2/K_3}) &= \mathrm{E}\left[\sum_{p=0}^{\infty}\frac{(X^2/K_3)^p}{p!}\right] = 1 + \mathrm{E}\left[\sum_{p=1}^{\infty}\frac{(X^2/K_3)^p}{p!}\right] \\
&= 1 + \mathrm{E}\left[\sum_{p=1}^{\infty}\frac{X^{2p}/K_3^p}{p!}\right] = 1 + \sum_{p=1}^{\infty}\frac{\mathrm{E}[X^{2p}]/K_3^p}{p!}.
\end{aligned}$$

By (c), $\mathrm{E}(X^{2p}) \leq K_2^{2p}(2p)^p$ for every $p \geq 1$. In addition, by Stirling's approximation,

$$p! \geq \sqrt{2\pi}p^{p+1/2}e^{-p} \geq p^p e^{-p} = \left(\frac{p}{e}\right)^p$$

for every $p \geq 1$. Thus,

$$\begin{aligned}
\mathrm{E}(e^{X^2/K_3}) &= 1 + \sum_{p=1}^{\infty}\frac{\mathrm{E}[X^{2p}]/K_3^p}{p!} \leq 1 + \sum_{p=1}^{\infty}\frac{K_2^{2p}(2p)^p/K_3^p}{(p/e)^p} \\
&= 1 + \sum_{p=1}^{\infty}\frac{K_2^{2p}(2p)^p/(4eK_2^2)^p}{(p/e)^p} = 1 + \sum_{p=1}^{\infty}\frac{(2p)^p/(4e)^p}{(p/e)^p} \\
&= 1 + \sum_{p=1}^{\infty}\left(\frac{1}{2}\right)^p = \sum_{p=0}^{\infty}\left(\frac{1}{2}\right)^p = 2.
\end{aligned}$$

(d) $\implies$ (b). Let $t > 0$. Then since $\mathrm{E}(e^{X^2/K_3}) \leq 2$ by (d), an application of Markov's Inequality yields

$$\mathrm{P}\{|X| > t\} = \mathrm{P}\{e^{X^2/K_3} > e^{t^2/K_3}\} \leq e^{-t^2/K_3}\mathrm{E}(e^{X^2/K_3}) \leq 2e^{-t^2/K_3}.$$

Thus, taking $K_1 = K_3$, we have

$$\mathrm{P}\{|X| > t\} \leq 2e^{-t^2/K_1}.$$

(c) $\implies$ (a). Let $p \geq 1$. Then since

$$[\mathrm{E}(|X|^p)]^{1/p} \leq K_2\sqrt{p}$$

by (c), an application of Minkowski's Inequality yields

$$\begin{aligned}
[\mathrm{E}(|X-\mu|^p)]^{1/p} &\leq [\mathrm{E}(|X|^p)]^{1/p} + [\mathrm{E}(|-\mu|^p)]^{1/p} = [\mathrm{E}(|X|^p)]^{1/p} + |\mu| \\
&\leq K_2\sqrt{p} + |\mu| \leq K_2\sqrt{p} + |\mu|\sqrt{p} \\
&= (K_2 + |\mu|)\sqrt{p} = K_2'\sqrt{p},
\end{aligned}$$

where $K_2' = K_2 + |\mu| > 0$. Thus, since $\mathrm{E}(X-\mu) = 0$, the result proven in the lectures implies that

$$\mathrm{E}(e^{\lambda(X-\mu)}) \le e^{\lambda^2/K_0} \quad \text{for any } \lambda \in \mathbb{R}$$

for some $K_0 > 0$.

(a) $\implies$ (c). Let $p \ge 1$. Since $\mathrm{E}(X-\mu) = 0$, the result proven in the lectures implies that

$$[\mathrm{E}(|X-\mu|^p)]^{1/p} \le K_2'\sqrt{p}$$

for some $K_2' > 0$. Thus by Minkowski's Inequality,

$$\begin{aligned}
[\mathrm{E}(|X|^p)]^{1/p} &= [\mathrm{E}(|(X-\mu)+\mu|^p)]^{1/p} \le [\mathrm{E}(|X-\mu|^p)]^{1/p} + [\mathrm{E}(|\mu|^p)]^{1/p} \\
&= [\mathrm{E}(|X-\mu|^p)]^{1/p} + |\mu| \le K_2'\sqrt{p} + |\mu| \\
&\le K_2'\sqrt{p} + |\mu|\sqrt{p} = (K_2' + \mu)\sqrt{p} = K_2\sqrt{p},
\end{aligned}$$

where $K_2 = K_2' + |\mu| > 0$.

4. Let $S \subseteq \{1, 2, \ldots, \beta\}$, and let $\nu_0 > 0$ and $\xi > 1$. Assume the restricted eigenvalue condition holds; that is, for any $b \in \mathbb{R}^p$ satisfying

$$\sum_{j \notin S} |b_j| \le \xi \sum_{j \in S} |b_j|, \tag{1}$$

we have

$$\nu_0^2 \left( \sum_{j \in S} b_j^2 \right) \le b^T \widetilde{\Sigma} b. \tag{2}$$

So let $b \in \mathbb{R}^p$ satisfy (1). Then $b$ satisfies (2). Then by Cauchy-Schwarz,

$$\left( \sum_{j \in S} |b_j| \right)^2 \le \left( \sum_{j \in S} b_j^2 \right) |S|.$$

Thus, by (2),

$$\nu_0^2 \left( \sum_{j \in S} |b_j| \right)^2 \le \nu_0^2 \left( \sum_{j \in S} b_j^2 \right) |S| \le (b^T \widetilde{\Sigma} b)|S|.$$

We therefore see that the compatibility condition holds. Hence the restricted eigenvalue condition implies the compatibility condition.

5. By definition, $\hat{\beta}$ minimizes the Lasso objective function. Thus, for any $0 \le t \le 1$,

$$\frac{1}{2}\|Y - X\hat{\beta}\|_n^2 + \lambda\|\beta\|_1 \le \frac{1}{2}\|Y - X[(1-t)\hat{\beta} + t\beta^*]\|_n^2 + \lambda\|(1-t)\hat{\beta} + t\beta^*\|_1. \tag{1}$$

Note that we have

$$\begin{aligned}
&\frac{1}{2}\|Y - X[(1-t)\hat{\beta} + t\beta^*]\|_n^2 \\
&\quad = \frac{1}{2}\|Y - X\hat{\beta} - tX(\hat{\beta} - \beta^*)\|_n^2 \\
&\quad = \frac{1}{2}\|Y - X\hat{\beta}\|_n^2 + \frac{1}{2}t^2\|X(\hat{\beta} - \beta^*)\|_n^2 - t\langle Y - X\hat{\beta}, X(\hat{\beta} - \beta^*)\rangle_n.
\end{aligned}$$

Thus we can write (1) as

$$\lambda\|\hat{\beta}\|_1 \leq \frac{1}{2}t^2\|X(\hat{\beta} - \beta^*)\|_n^2 - t\langle Y - X\hat{\beta}, X(\hat{\beta} - \beta^*\rangle_n + \lambda\|(1-t)\hat{\beta} + t\beta^*)\|_1.$$

[This is as far as I got – all attempts hereafter were fruitless. Is there a trick, or is this just a lot of elaborate algebra?]