Another version of Bernstein's inequality

December 16, 2017

The purpose of this note is to remove one of the boundedness assumptions in Bernstein's inequality as stated in [1], Theorem 3. All undefined notation is taken from [1].

Let $f:\Omega=\prod_{i=1}^{n}\Omega_i\rightarrow\mathbb{R}$ and consider the three conditions

$$\begin{aligned}
(A) &= \left((f-E_kf)\leq b \text{ for all } k\right)\\
(B) &= \left(E_k\left[(f-E_kf)^m\right]\leq\frac{1}{2}m!\sigma_k^2(f)\,b^{m-2}\text{ for } m\geq 2\text{ and all } k\right)\\
(C) &= \left(\sum_{k=1}^{n}E_k\left[(f-E_kf)^m\right]\leq\frac{\Sigma^2(f)}{2}m!b^{m-2}\text{ for } m\geq 2\right)
\end{aligned}$$

Then $(A)\implies(B)\implies(C)$. The last condition is sufficient for the following version of Bernstein's inequality, which extends Theorem 2.10 in [2] from sums to general functions.

Theorem 1 *Let $f:\Omega=\prod_{i=1}^{n}\Omega_i\rightarrow\mathbb{R}$ be measurable and suppose that (C) holds. Then for $t>0$*

$$\Pr\{f-Ef>t\}\leq\exp\left(\frac{-t^2}{2E\left[\Sigma^2(f)\right]+(2b+J_\mu)\,t}\right).$$

By rescaling it suffices to prove this for $b=1$. The next Lemma replaces Lemma 9 in [1].

Lemma 2 *Suppose (C) holds with $b=1$. Then for all $\beta\geq 0$*

$$S_f(\beta)\leq\frac{\beta^2E_{\beta f}\left[\Sigma^2(f)\right]}{2(1-\beta)^2}.$$

Proof. First we get from the variational property of variance, that

$$\begin{aligned}
\sigma_{k,\beta f}^2(f) &\leq E_{k,\beta f}\left[(f-E_k(f))^2\right]=\frac{E_k\left[(f-E_k(f))^2e^{\beta(f-Ef)}\right]}{E_k\left[e^{\beta(f-Ef)}\right]}\\
&\leq E_k\left[(f-E_k(f))^2e^{\beta(f-Ef)}\right],
\end{aligned}$$

where we used Jensen's inequality to get $E_k\left[\exp\left(\beta\left(f-Ef\right)\right)\right]\geq 1$ for the second inequality. From monotone convergence and (C) we then get

$$\begin{aligned}\sum_{k=1}^{n}\sigma_{k,\beta f}^{2}\left(f\right) &\leq \sum_{k=1}^{n}E_{k}\left[\left(f-E_{k}f\right)^{2}e^{\beta\left(f-E_{k}f\right)}\right]=\sum_{n=0}^{\infty}\sum_{k=1}^{n}\frac{\beta^{m}}{m!}E\left[\left(f-E_{k}f\right)^{m+2}\right] \\ &\leq \frac{\Sigma^{2}\left(f\right)}{2}\sum_{m=0}^{\infty}\left(m+1\right)\left(m+2\right)\beta^{m}.\end{aligned}$$

Thus

$$\begin{aligned}S_{f}\left(\beta\right) &\leq E_{\beta f}\left[\int_{0}^{\beta}\int_{t}^{\beta}\sum_{k=1}^{n}\sigma_{k,\beta f}^{2}\left(f\right)\ ds\ dt\right] \\ &\leq \frac{E_{\beta f}\left[\Sigma^{2}\left(f\right)\right]}{2}\sum_{m=0}^{\infty}\left(m+1\right)\left(m+2\right)\int_{0}^{\beta}\int_{t}^{\beta}s^{m}dsdt \\ &= \frac{E_{\beta f}\left[\Sigma^{2}\left(f\right)\right]}{2}\beta^{2}\sum_{m=0}^{\infty}\left(m+1\right)\beta^{m}=\frac{\beta^{2}E_{\beta f}\left[\Sigma^{2}\left(f\right)\right]}{2\left(1-\beta\right)^{2}}.\end{aligned}$$

■

The next proposition replaces Proposition 16 in [1].

Proposition 3 *Suppose that $f:\Omega\to\mathbb{R}$ is such that (C) holds with $b=1$, and that*

$$D\left(\Sigma^{2}\left(f\right)\right)\leq a^{2}\ \Sigma^{2}\left(f\right),$$

with $a\geq 0$. Then for all $t>0$

$$\Pr\left\{f-Ef>t\right\}\leq\exp\left(\frac{-t^{2}}{2E\left[\Sigma^{2}\left(f\right)\right]+\left(2+a\right)t}\right).$$

Proof. We can assume $a>0$. Let $0<\gamma\leq\beta<1/\left(1+a/2\right)$ and set $\theta=\gamma/\left(a\left(1-\gamma\right)\right)$. Then $\gamma^{2}/\left(2\left(1-\gamma\right)^{2}\right)<\theta<2/a^{2}$. By the Lemma 2

$$\theta S_{f}\left(\gamma\right)\leq\frac{\gamma^{2}}{2\left(1-\gamma\right)^{2}}E_{\gamma f}\left[\theta\Sigma^{2}\left(f\right)\right]\leq\frac{\gamma^{2}}{2\left(1-\gamma\right)^{2}}\left(S_{f}\left(\gamma\right)+\ln E\left[e^{\theta\Sigma^{2}\left(f\right)}\right]\right),$$

where the second inequality follows from the decoupling lemma (Lemma 13 in [1]). Subtract $\gamma^{2}/\left(2\left(1-\gamma\right)^{2}\right)S_{f}\left(\gamma\right)$ to get

$$S_{f}\left(\gamma\right)\left(\theta-\frac{\gamma^{2}}{2\left(1-\gamma\right)^{2}}\right)\leq\frac{\gamma^{2}}{2\left(1-\gamma\right)^{2}}\ln E\left[e^{\theta\Sigma^{2}\left(f\right)}\right].$$

Since $\gamma^{2}/\left(2\left(1-\gamma\right)^{2}\right)<\theta$ this simplifies, using the value of θ, to

$$S_{f}\left(\gamma\right)\leq\frac{\gamma a}{2\left(1-\left(1+\left(a/2\right)\right)\gamma\right)}\ln E\left[e^{\theta\Sigma^{2}\left(f\right)}\right]. \tag{1}$$

On the other hand $\theta < 2/a^2$, so the assumed self-boundedness of $\Sigma^2(f)$ and Lemma 12 in [1] give

$$\ln E\left[e^{\theta\Sigma^2(f)}\right] \leq \frac{\theta}{1-a^2\theta/2} E\left[\Sigma^2(f)\right] = \frac{\gamma/a}{1-(1+a/2)\gamma} E\left[\Sigma^2(f)\right]. \tag{2}$$

Combining (1) and (2) to get a bound on $S_f(\gamma)$ gives

$$\int_0^\beta \frac{S_f(\gamma)\,d\gamma}{\gamma^2} \leq E\left[\Sigma^2(f)\right] \int_0^\beta \frac{d\gamma}{2\left(1-(1+a/2)\gamma\right)^2} = \frac{E\left[\Sigma^2(f)\right]}{2} \frac{\beta}{1-(1+a/2)\beta}$$

and from Lemma 8 in [1] and Lemma 15 in [1]

$$\begin{aligned}
\Pr\{f - Ef > t\} &\leq \inf_{\beta>0} \exp\left(\beta \int_0^\beta \frac{S_f(\gamma)}{\gamma^2} d\gamma - \beta t\right) \\
&\leq \inf_{\beta\in(0,1/(1/3+a/2))} \exp\left(\frac{E\left[\Sigma^2(f)\right]}{2} \frac{\beta^2}{1-(1+a/2)\beta} - \beta t\right) \\
&\leq \exp\left(\frac{-t^2}{2\left(E\left[\Sigma^2(f)\right] + (1+a/2)t\right)}\right).
\end{aligned}$$

■

Theorem 1 now follows from combining Proposition 3 with Proposition 17 in [1].

References

[1] Andreas Maurer. "A Bernstein-type inequality for functions of bounded interaction." arXiv preprint arXiv:1701.06191 (2017).

[2] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. Concentration inequalities: A nonasymptotic theory of independence. Oxford university press, 2013.