A Self-normalized Martingale Tail Inequality

The self-normalized martingale tail inequality that we present here is the scalar-valued version of the more general vector-valued results obtained by Abbasi-Yadkori et al. (2011b,a). We include the proof for completeness.

Theorem 7 (Self-normalized bound for martingales)**.** *Let $\{F_t\}_{t=1}^{\infty}$ be a filtration. Let τ be a stopping time w.r.t. to the filtration $\{F_{t+1}\}_{t=1}^{\infty}$ i.e. the event $\{\tau \le t\}$ belongs to F_{t+1}. Let $\{Z_t\}_{t=1}^{\infty}$ be a sequence of real-valued variables such that Z_t is F_t-measurable. Let $\{\eta_t\}_{t=1}^{\infty}$ be a sequence of real-valued random variables such that η_t is F_{t+1}-measurable and is conditionally R-sub-Gaussian. Let $V > 0$ be deterministic. Then, for any $\delta > 0$, with probability at least $1-\delta$,*

$$\frac{\left(\sum_{t=1}^{\tau} \eta_t Z_t\right)^2}{V + \sum_{t=1}^{\tau} Z_t^2} \le 2R^2 \ln\left(\frac{\sqrt{V + \sum_{t=1}^{\tau} Z_t^2}}{\delta\sqrt{V}}\right) .$$

Proof. Pick $\lambda \in \mathbb{R}$ and let

$$D_t^\lambda = \exp\left(\frac{\eta_t \lambda Z_t}{R} - \frac{1}{2}\lambda^2 Z_t^2\right) ,$$

$$S_t = \sum_{s=1}^{t} \eta_s \lambda Z_s ,$$

$$M_t^\lambda = \exp\left(\frac{\lambda S_t}{R} - \frac{1}{2}\lambda^2 \sum_{s=1}^{t} Z_t^2\right) .$$

We claim that $\{M_t^\lambda\}_{t=1}^{\infty}$ is an $\{F_{t+1}\}_{t=1}^{\infty}$-adapted supermartingale. That $M_t^\lambda \in F_{t+1}$ for $t = 1, 2, \ldots$ is clear from the definitions. By sub-Gaussianity, $\mathbf{E}[D_t^\lambda \mid F_t] \le 1$. Further,

$$\begin{aligned} \mathbf{E}[M_t^\lambda | F_t] &= \mathbf{E}[M_{t-1}^\lambda D_t^\lambda \mid F_t] \\ &= M_{t-1}^\lambda \, \mathbf{E}[D_t^\lambda \mid F_t] \le M_{t-1}^\lambda , \end{aligned}$$

showing that $\{M_t\}_{t=1}^{\infty}$ is indeed a supermartingale.

Next we show that M_τ^λ is always well-defined and $\mathbf{E}[M_\tau^\lambda] \le 1$. First define $\tilde{M} = M_\tau^\lambda$ and note that $\tilde{M}(\omega) = M_{\tau(\omega)}^\lambda(\omega)$. Thus, when $\tau(\omega) = \infty$, we need to argue about $M_\infty^\lambda(\omega)$. By the convergence theorem for nonnegative supermartingales, $\lim_{t\to\infty} M_t^\lambda(\omega)$ is well-defined, which means M_τ^λ is well-defined, independently of whether $\tau < \infty$ holds or not. Now let $Q_t^\lambda = M_{\min\{\tau,t\}}^\lambda$ be a stopped version of M_t^λ. We proceed by using Fatou's Lemma to show that $\mathbf{E}[M_\tau^\lambda] = \mathbf{E}[\liminf_{t\to\infty} Q_t^\lambda] \le \liminf_{t\to\infty} \mathbf{E}[Q_t^\lambda] \le 1$.

Let F_∞ be the σ-algebra generated by $\{F_t\}_{t=1}^{\infty}$ i.e. the tail σ-algebra. Let Λ be a zero-mean Gaussian random variable with variance $1/V$ independent of F_∞. Define $M_t = \mathbf{E}[M_t^\Lambda \mid F_\infty]$. Clearly, we still have $\mathbf{E}[M_\tau] = \mathbf{E}[M_\tau^\Lambda] = \mathbf{E}[\mathbf{E}[M_\tau^\Lambda] \mid \Lambda] \le \mathbf{E}[1 \mid \Lambda] \le 1$.

Let us calculate M_t. We will need the density λ which is $f(\lambda) = \frac{1}{\sqrt{2\pi/V}} e^{-V\lambda^2/2}$. Now, it is easy to write M_t explicitly

$$\begin{aligned} M_t &= \mathbf{E}[M_t^\Lambda \mid F_\infty] \\ &= \int_{-\infty}^{\infty} M_t^\lambda f(\lambda) \, d\lambda \\ &= \sqrt{\frac{V}{2\pi}} \int_{\infty}^{\infty} \exp\left(\frac{\lambda S_t}{R} - \frac{\lambda^2}{2}\sum_{s=1}^{t} Z_t^2\right) e^{-V\lambda^2/2} \, d\lambda \\ &= \exp\left(\frac{S_t^2}{2R^2(V + \sum_{s=1}^{t} Z_t^2)}\right) \sqrt{\frac{V}{V + \sum_{s=1}^{t} Z_t^2}} , \end{aligned}$$

where we have used that $\int_{-\infty}^{\infty} \exp(a\lambda - b\lambda^2) = \exp(a^2/(4b))\sqrt{\pi/b}$.

To finish the proof, we use Markov's inequality and the fact that $\mathbf{E}[M_\tau] \leq 1$:

$$\begin{aligned}
&\Pr\left[\frac{\left(\sum_{t=1}^{\tau}\eta_t Z_t\right)^2}{V+\sum_{t=1}^{\tau} Z_t^2} \geq 2R^2 \ln\left(\frac{\sqrt{V+\sum_{t=1} Z_t^2}}{\delta\sqrt{V}}\right)\right] \\
&= \Pr\left[\frac{S_\tau^2}{2R^2(V+\sum_{t=1}^{\tau} Z_t^2)} \geq \ln\left(\frac{\sqrt{V+\sum_{t=1} Z_t^2}}{\delta\sqrt{V}}\right)\right] \\
&= \Pr\left[\exp\left(\frac{S_\tau^2}{2R^2(V+\sum_{t=1}^{\tau} Z_t^2)}\right) \geq \frac{\sqrt{V+\sum_{t=1} Z_t^2}}{\delta\sqrt{V}}\right] \\
&= \Pr\left[M_\tau \geq \frac{1}{\delta}\right] \\
&\leq \delta\ .
\end{aligned}$$

□

The theorem can be "bootstrapped" to a "stronger" statement (or at least one, that looks stronger at the first sight) that holds uniformly for all time steps t as opposed to only a particular (stopping) time τ. The idea of the proof goes back at least to Freedman (1975).

Corollary 8 (Uniform Bound)**.** *Under the same assumptions as the previous theorem, for any $\delta > 0$, with probability at least $1-\delta$, for all $n \geq 0$,*

$$\left|\sum_{t=1}^{n}\eta_t Z_t\right| \leq R\sqrt{2\left(V+\sum_{t=1}^{n} Z_t^2\right)\ln\left(\frac{\sqrt{V+\sum_{t=1}^{n} Z_t^2}}{\delta\sqrt{V}}\right)}\ .$$

Proof. Define the "bad" event

$$B_t(\delta) = \left\{\omega\in\Omega\ :\ \frac{\left(\sum_{s=1}^{t}\eta_s Z_s\right)^2}{V+\sum_{s=1}^{t} Z_s^2} > 2R^2\ln\left(\frac{\sqrt{V+\sum_{s=1}^{t} Z_s^2}}{\delta\sqrt{V}}\right)\right\}.$$

We are interested in bounding the probability that $\bigcup_{t\geq 0} B_t(\delta)$ happens. Define $\tau(\omega) = \min\{t\geq 0\ :\ \omega\in B_t(\delta)\}$, with the convention that $\min\emptyset = \infty$. Then, τ is a stopping time. Further,

$$\bigcup_{t\geq 0} B_t(\delta) = \{\omega\ :\ \tau(\omega) < \infty\}.$$

Thus, by Theorem 7 it holds that

$$\begin{aligned}
\Pr\left[\bigcup_{t\geq 0} B_t(\delta)\right] &= \Pr\left[\tau<\infty\right] \\
&= \Pr\left[\frac{\left(\sum_{t=1}^{\tau}\eta_t Z_t\right)^2}{V+\sum_{t=1}^{\tau} Z_t^2} > 2R^2\ln\left(\frac{\sqrt{V+\sum_{t=1} Z_t^2}}{\delta\sqrt{V}}\right) \text{ and } \tau<\infty\right] \\
&= \Pr\left[\frac{\left(\sum_{t=1}^{\tau}\eta_t Z_t\right)^2}{V+\sum_{t=1}^{\tau} Z_t^2} > 2R^2\ln\left(\frac{\sqrt{V+\sum_{t=1} Z_t^2}}{\delta\sqrt{V}}\right)\right] \\
&\leq \delta\ .
\end{aligned}$$

□

B Some Useful Tricks

Proposition 9 (Square-Root Trick)**.** *Let $a,b\geq 0$. If $z^2 \leq a+bz$ then $z\leq b+\sqrt{a}$.*

Proof of the Proposition 9. Let $q(x) = x^2 - bx - a$. The condition $z^2 \le a + bz$ can be expressed as $q(z) \le 0$. The quadratic polynomial $q(x)$ has two roots

$$x_{1,2} = \frac{b \pm \sqrt{b^2 + 4a}}{2} \, .$$

The condition $q(z) \le 0$ implies that $z \le \max\{x_1, x_2\}$. Therefore,

$$z \le \max\{x_1, x_2\} = \frac{b + \sqrt{b^2 + 4a}}{2} \le b + \sqrt{a} \, ,$$

where we have used that $\sqrt{u+v} \le \sqrt{u} + \sqrt{v}$ holds for any $u, v \ge 0$. □

Proposition 10 (Logarithmic Trick)**.** *Let $c \ge 1$, $f > 0$, $\delta \in (0, 1/4]$. If $z \ge 1$ and $z \le c + f\sqrt{\ln(z/\delta)}$ then $z \le c + f\sqrt{2\ln\left(\frac{c+f}{\delta}\right)}$.*

Proof of the Proposition 10. Let $g(x) = x - c - f\sqrt{\ln(x/\delta)}$ for any $x \ge 1$. The condition $z \le c + f\sqrt{\ln(z/\delta)}$ can be expressed as $g(z) \le 0$. For large enough x, the function $g(x)$ is increasing. This is easy to see, since $g'(x) = 1 - \frac{f}{2x\sqrt{\ln(x/\delta)}}$. Namely, it is not hard see $g(x)$ is increasing for $x \ge \max\{1, f/2\}$ since for any such x, $g'(x)$ is positive.

Clearly, $c + f\sqrt{2\ln\left(\frac{c+f}{\delta}\right)} \ge \max\{1, f/2\}$ since $c \ge 1$ and $\delta \in (0, 1/4]$. Therefore, it suffices to show that

$$g\left(c + f\sqrt{2\ln\left(\frac{c+f}{\delta}\right)}\right) \ge 0 \, .$$

This is verified by the following calculation

$$\begin{aligned}
g\left(c + f\sqrt{2\ln\left(\frac{c+f}{\delta}\right)}\right) &= c + f\sqrt{2\ln\left(\frac{c+f}{\delta}\right)} - c - f\sqrt{\ln\left(\frac{c + f\sqrt{2\ln\left((c+f)/\delta\right)}}{\delta}\right)} \\
&= f\sqrt{2\ln\left(\frac{c+f}{\delta}\right)} - f\sqrt{\ln\left(\frac{c + f\sqrt{2\ln\left((c+f)/\delta\right)}}{\delta}\right)} \\
&= f\sqrt{\ln\left(\frac{c+f}{\delta}\right)^2} - f\sqrt{\ln\left(\frac{c + f\sqrt{2\ln\left((c+f)/\delta\right)}}{\delta}\right)} \\
&\ge f\sqrt{\ln\left(\frac{c+f}{\delta}\right)^2} - f\sqrt{\ln\left(\frac{(c+f)\sqrt{2\ln\left((c+f)/\delta\right)}}{\delta}\right)} \\
&= f\sqrt{\ln\left(A^2\right)} - f\sqrt{\ln\left(A\sqrt{2\ln A}\right)} \\
&\ge 0,
\end{aligned}$$

where have defined $A = (c+f)/\delta$ and the last inequality follows from that $A^2 \ge A\sqrt{2\ln A}$ for any $A > 0$. □

C Proof of Theorem 3

In this section we will need the following notation. For a given positive definite matrix $A \in \mathbb{R}^{d \times d}$ we denote by $\langle x, y \rangle_A = x^\top A y$ the inner product between two vectors $x, y \in \mathbb{R}^d$ induced by A. We denote by $\|x\|_A = \sqrt{\langle x, x \rangle_A} = \sqrt{x^\top A x}$ the corresponding norm.

The following lemma is a from Dani et al. (2008). We reproduce the proof for completeness.

Lemma 11 (Elliptical Potential)**.** *Let $x_1, x_2, \ldots, x_n \in \mathbb{R}^d$ and let $V_t = I + \sum_{s=1}^t x_s^\top x_s$ for $t = 0, 1, 2, \ldots, n$. Then it holds that*

$$\sum_{t=1}^n \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \leq 2\ln(\det(V_n)) \ .$$

Furthermore, if $\|x_t\|_2 \leq X$ for all $t = 1, 2, \ldots, n$ then

$$\ln(\det(V_n)) \leq d \ln\left(1 + \frac{nX^2}{d}\right) \ .$$

Proof of Lemma 11 . We use the inequality $x \leq 2\ln(1+x)$ valid for all $x \in [0,1]$:

$$\sum_{t=1}^n \min\left\{1, \|x_t\|^2_{V_{t-1}^{-1}}\right\} \leq \sum_{t=1}^n 2\ln\left(1 + \|x_t\|^2_{V_{t-1}^{-1}}\right) = 2\ln\left(\prod_{t=1}^n \left(1 + \|x_t\|^2_{V_{t-1}^{-1}}\right)\right) \ .$$

We now show that $\det(V_n) = \prod_{t=1}^n (1 + \|x_t\|^2_{V_{t-1}^{-1}})$:

$$\begin{aligned}
\det(V_n) &= \det(V_{n-1} + x_n x_n^\top) \\
&= \det\left(V_{n-1}(I + (V_{n-1}^{-1/2} x_n)(V_{n-1}^{-1/2} x_n)^\top\right) \\
&= \det(V_{n-1}) \det\left(I + (V_{n-1}^{-1/2} x_n)(V_n^{-1/2} x_n)^\top\right) \\
&= \det(V_{n-1}) \cdot \left(1 + \|x_n\|^2_{V_{n-1}^{-1}}\right) \\
&= \cdots \\
&= \prod_{t=1}^n (1 + \|x_t\|^2_{V_{t-1}^{-1}}) \ . && \text{(since } V_0 = I\text{)}
\end{aligned}$$

In the above calculation we have used that $\det(I + zz^\top) = 1 + \|z\|_2^2$ since all but one eigenvalue of $I + zz^\top$ equals to 1 and the remaining eigenvalue is $1 + \|z\|_2^2$ with associated eigenvector z.

To prove the second part, consider the eigenvalues $\alpha_1, \alpha_2, \ldots, \alpha_d$ of V_n. Since V_n is positive definite, the eigenvalues are positive. Recall that $\det(V_n) = \prod_{i=1}^d \alpha_i$. The bound on $\|x_t\| \leq X$ implies a bound on the trace of V_n:

$$\operatorname{Trace} V_n = \operatorname{Trace}(I) + \sum_{t=1}^n \operatorname{Trace}(x_t x_t^\top) = d + \sum_{t=1}^n \|x_t\|_2^2 \leq d + nX^2 \ .$$

Recalling that $\operatorname{Trace}(V_n) = \sum_{i=1}^d \alpha_i$ we can apply the AM-GM inequality:

$$\sqrt[d]{\alpha_1 \alpha_2 \cdots \alpha_d} \leq \frac{\alpha_1 + \alpha_2 + \cdots + \alpha_d}{d} = \frac{\operatorname{Trace}(V_n)}{d} \ ,$$

from which the second inequality follows by taking logarithm and multiplying by d. ☐

Proof of Theorem 3. Consider the event A when $\theta_* \in \bigcap_{t=0}^\infty C_t$. By Corollary 2, the event A occurs with probability at least $1 - \delta$.

The set C_{t-1} is an ellipsoid underlying the covariance matrix $V_{t-1} = I + \sum_{s=1}^{t-1} X_s^\top X_s$ and center

$$\widehat{\theta}_t = \operatorname*{argmin}_{\theta \in \mathbb{R}^d} \left(\|\theta\|_2^2 + \sum_{s=1}^{t-1} (\widehat{Y}_s - \langle \theta, X_s \rangle)^2\right) \ .$$

The ellipsoid C_{t-1} is non-empty since θ_* lies in it (on the event A). Therefore $\widehat{\theta}_t \in C_{t-1}$. We can thus express the ellipsoid as

$$C_{t-1} = \left\{\theta \in \mathbb{R}^d \ : \ (\theta - \widehat{\theta}_t)^\top V_{t-1} (\theta - \widehat{\theta}_t) + \|\widehat{\theta}_t\|_2^2 + \sum_{s=1}^{t-1} \left(\widehat{Y}_s - \left\langle \widehat{\theta}_t, X_s \right\rangle\right)^2 \leq \beta_{t-1}(\delta)\right\}.$$

The ellipsoid is contained in a larger ellipsoid

$$C_{t-1} \subseteq \left\{ \theta \in \mathbb{R}^d \ : \ (\theta - \widehat{\theta}_t)^\top V_{t-1} (\theta - \widehat{\theta}_t) \le \beta_{t-1}(\delta) \right\} = \left\{ \theta \in \mathbb{R}^d \ : \ \|\theta - \widehat{\theta}_t\|_{V_{t-1}} \le \sqrt{\beta_{t-1}(\delta)} \right\} .$$

First, we bound the instantaneous regret using that $(X_t, \widetilde{\theta}_t) = \operatorname{argmax}_{(x,\theta) \in D_t \times C_{t-1}} \langle x, \theta \rangle$:

$$\begin{aligned}
\langle x_* - X_t, \theta_* \rangle &= \langle x_*, \theta_* \rangle - \langle X_t, \theta_* \rangle \\
&\le \left\langle X_t, \widetilde{\theta}_t \right\rangle - \langle X_t, \theta_* \rangle \\
&= \left\langle X_t, \widetilde{\theta}_t - \theta_* \right\rangle \\
&= \left\langle X_t, \widetilde{\theta}_t - \widehat{\theta}_t \right\rangle - \left\langle X_t, \widehat{\theta}_t - \theta_* \right\rangle \\
&\le \left| \left\langle X_t, \widetilde{\theta}_t - \widehat{\theta}_t \right\rangle \right| + \left| \left\langle X_t, \widehat{\theta}_t - \theta_* \right\rangle \right| \\
&\le \|X_t\|_{V_{t-1}^{-1}} \left\| \widetilde{\theta}_t - \widehat{\theta}_t \right\|_{V_{t-1}} + \|X_t\|_{V_{t-1}^{-1}} \left\| \widehat{\theta}_t - \theta_* \right\|_{V_{t-1}} && \text{(Cauchy-Schwarz)} \\
&\le 2\sqrt{\beta_{t-1}(\delta)} \cdot \|X_t\|_{V_{t-1}^{-1}} . && \text{(because } \widetilde{\theta}_t, \theta_* \in C_{t-1}\text{)}
\end{aligned}$$

Since we assume that $|\langle x, \theta_* \rangle| \le G$ for any $x \in D_t$ and any $t = 1, 2, \ldots, n$, we can upper bound $\langle x_* - X_t, \theta_* \rangle \le 2 \min\{G, \sqrt{\beta_{t-1}(\delta)} \cdot \|X_t\|_{V_{t-1}^{-1}}\}$. Summing over all t we upper bound regret

$$\begin{aligned}
R_n &= \sum_{t=1}^n \langle x^* - X_t, \theta_* \rangle \\
&\le 2 \sum_{t=1}^n \min \left\{ G, \sqrt{\beta_{t-1}(\delta)} \cdot \|X_t\|_{V_{t-1}^{-1}} \right\} \\
&\le 2 \sum_{t=1}^n \sqrt{\beta_{t-1}(\delta)} \cdot \min \left\{ G, \|X_t\|_{V_{t-1}^{-1}} \right\} && \text{(since } \beta_{t-1}(\delta) \ge 1\text{)} \\
&\le 2 \left(\max_{0 \le t < n} \sqrt{\beta_t(\delta)} \right) \sum_{t=1}^n \min \left\{ G, \|X_t\|_{V_{t-1}^{-1}} \right\} \\
&\le 2 \left(\max_{0 \le t < n} \sqrt{\beta_t(\delta)} \right) \max\{1, G\} \sum_{t=1}^n \min \left\{ 1, \|X_t\|_{V_{t-1}^{-1}} \right\} \\
&\le 2 \left(\max_{0 \le t < n} \sqrt{\beta_t(\delta)} \right) \max\{1, G\} \times \sqrt{ n \sum_{t=1}^n \min \left\{ 1, \|X_t\|_{V_{t-1}^{-1}}^2 \right\} } && \text{(Cauchy-Schwarz)} \\
&\le 2 \max\{1, G\} \sqrt{ 2nd \log \left(1 + \frac{nX^2}{d} \right) \max_{0 \le t < n} \beta_t(\delta) } ,
\end{aligned}$$

where the last inequality follows from Lemma 11. □

Proof of Theorem 4. Summing over all t we upper bound regret

$$R_n = \sum_{t=1}^n \langle x^* - X_t, \theta_* \rangle \le \frac{1}{\Delta} \sum_{t=1}^n \langle x^* - X_t, \theta_* \rangle^2 ,$$

where the last inequality follows from the fact that either $\langle x^* - X_t, \theta_* \rangle = 0$ or $\langle x^* - X_t, \theta_* \rangle > \Delta$. Then we take similar steps as in the proof of Theorem 3 to obtain

$$\begin{aligned} R_n &\le \frac{1}{\Delta} \sum_{t=1}^{n} \langle x^* - X_t, \theta_* \rangle^2 \\ &\le \frac{4}{\Delta} \left(\max_{0 \le t < n} \beta_t(\delta) \right) \max\{1, G^2\} \sum_{t=1}^{n} \min\left\{1, \|X_t\|^2_{V_{t-1}^{-1}} \right\} \\ &\le \frac{8d}{\Delta} \left(\max_{0 \le t < n} \beta_t(\delta) \right) \max\{1, G^2\} \log\left(1 + \frac{nX^2}{d}\right), \end{aligned}$$

finishing the proof of the problem dependent bound. □