Moments of sample covariance and precision matrices

T. Mitchell Roddenberry

11 August 2024

1 Introduction

For a random sample covariance matrix C from n samples of a probability distribution on $\mathbb{R}^m$, we consider the distribution of its extreme eigenvalues. In particular, we are concerned with the moments of the distribution of eigenvalues. We refer to the moments of the maximum eigenvalue as the nonnegative moments of the matrix, *i.e.*, for $q \geq 0$, the expected value $\mathbb{E}[\|C^q\|]$. Similarly, we refer to the moments of the inverse minimimum eigenvalue as the negative moments of the matrix, *i.e.*, for $q \geq 0$, the expected value $\mathbb{E}[\|C^{-q}\|]$. We are interested in the easy regime of $n \to \infty$.

In Section 2, we will show that as $n \to \infty$, the nonnegative moments satisfy $\mathbb{E}[\|C^q\|] = O(1)$ for subgaussian distributions. Subgaussianity is not strong enough to control the negative moments, so in Section 3 we show that in the special case of a gaussian distribution, the sample covariance satisfies $\mathbb{E}[\|C^{-q}\|] = O(m^{1+\epsilon})$ for arbitrarily small $\epsilon > 0$.

2 Nonnegative Moments of a Random Matrix

Let X be an $m \times n$ matrix with independent, mean-zero, subgaussian entries. Then, for any $t > 0$, we have by [Ver18, Theorem 4.4.5],

$$\|X\| \geq C_0 K \left(\sqrt{m} + \sqrt{n} + t\right) \tag{1}$$

with probability less that $2\exp(-t^2)$, where $C_0 > 0$ is some absolute constant, and K is the maximum subgaussian norm of the entries of X. Let us assume that $K = 1$, for simplicity. For $C = n^{-1}XX^\top$, we can bound the nonnegative moments of $\|C\| = n^{-1}\|X\|^2$ by integrating this tail bound. Let $q \geq 1$, and consider

$$\begin{aligned}
\mathbb{E}[\|C^q\|] &= \mathbb{E}[\|C\|^q] \\
&= \int_0^\infty \mathbb{P}\left(\|C\|^q \geq t\right) dt \\
&= 2q \cdot n^{-q} \int_0^\infty \mathbb{P}\left(\|X\| \geq u\right) u^{2q-1} du \\
&\leq C_0^{2q-1} \left(\left(\frac{\sqrt{n} + \sqrt{m} + \sqrt{\ln(2)}}{\sqrt{n}} \right)^{2q} + 4q \int_{\sqrt{\ln(2)}}^\infty e^{-t^2} \frac{(\sqrt{n} + \sqrt{m} + t)^{2q-1}}{n^q} dt \right).
\end{aligned} \tag{2}$$

As $n \to \infty$, it is clear that this converges to a constant, indicating that $\mathbb{E}[\|C^q\|] = O_n(1)$.

3 Negative Moments of a Random Matrix

We now move on to the moments of the matrix inverse. The assumption of subgaussianity used in Section 2 is not sufficient here. To see why, consider the Rademacher random variable, which takes values $\{-1, 1\}$ with equal probability. Taking X to be a random matrix with independent Rademacher-valued entries, it is clear that X is singular with nonzero (albeit, tending to zero as $n \to \infty$) probability. That is to say, $\|(XX^\top)^{-q}\| = \infty$ with nonzero probability, which obstructs the existence of the q^{th} moment.

We now consider the guassian case, so that $XX^\top$ follows a Wishart distribution, which is almost surely nonsingular. For any $u \in \mathbb{R}^m$, by [Eat07, Proposition 8.9], the random variable $\langle u, (XX^\top)^{-1}u\rangle$ follows an inverse-χ^2 distribution with $n - m + 1$ degrees of freedom. Owing to the fact that $(XX^\top)^{-1}$ is positive semidefinite, it also holds that for any orthonormal basis $\{u_j\}_{j=1}^m$ for $\mathbb{R}^m$, we have

$$\max_j \langle u_j, (XX^\top)^{-1}u_j\rangle \leq \|(XX^\top)^{-1}\| \leq m \cdot \max_j \langle u_j, (XX^\top)^{-1}u_j\rangle. \tag{3}$$

The left-hand inequality is obvious, and the right-hand inequality follows by considering the trace of $(XX^\top)^{-1}$. Recalling the definition of the sample covariance as $C = n^{-1}XX^\top$, we bound the negative moments of C via a union bound argument over an orthonormal basis. Let $q \geq 1$, and let an orthonormal basis $\{u_j\}_{j=1}^m$ for $\mathbb{R}^m$ be given arbitrarily. Then, we have

$$\|C^{-1}\|^q \leq m^q \cdot \left(\max_j \langle u_j, C^{-1}u_j\rangle\right)^q, \tag{4}$$

which implies via a union bound argument, for any $t > 0$,

$$\begin{aligned}\mathbb{P}\left(\|C^{-1}\|^q \geq t\right) &\leq \mathbb{P}\left(\max_j \langle u_j, C^{-1}u_j\rangle \geq t^{1/q}/m\right) \\ &\leq \sum_{j=1}^m \mathbb{P}\left(\langle u_j, C^{-1}u_j\rangle \geq t^{1/q}/m\right).\end{aligned} \tag{5}$$

We can now bound the moment by integration.

$$\begin{aligned}\mathbb{E}[\|C^{-q}\|] &= \mathbb{E}[\|C^{-1}\|^q] \\ &= \int_0^\infty \mathbb{P}\left(\|C^{-1}\|^q \geq t\right) dt \\ &\leq \sum_{j=1}^m \int_0^\infty \mathbb{P}\left(\langle u_j, C^{-1}u_j\rangle \geq t^{1/q}/m\right) dt \\ &= \sum_{j=1}^m \int_0^\infty \mathbb{P}\left(\langle u_j, (XX^\top)^{-1}u_j\rangle \geq t^{1/q}/(mn)\right) dt.\end{aligned} \tag{6}$$

Each summand is merely (a scaled version of) the q^{th} moment of an inverse-χ^2 random variable with $n - m + 1$ degrees of freedom, which is known to be

$$\int_0^\infty \mathbb{P}\left(\langle u_j, (XX^\top)^{-1}u_j\rangle \geq t^{1/q}/(mn)\right) dt = \left(\frac{m \cdot n}{2}\right)^q \frac{\Gamma\left(\frac{n-m+1}{2} - q\right)}{\Gamma\left(\frac{n-m+1}{2}\right)}. \tag{7}$$

As $n \to \infty$, this clearly tends to m^q, so that $\mathbb{E}[\|C^{-q}\|] = O_n(m^{q+1})$. Since this argument holds for arbitrary q, Lyapunov's inequality can be applied to yield the tighter bound $\mathbb{E}[\|C^{-q}\|] = O_n(m^{q+\epsilon})$ for arbitrary $\epsilon > 0$. This is in contrast to the $O_n(1)$ bound on $\mathbb{E}[\|C^q\|]$, which is independent of m.

References

[Eat07] Morris L Eaton. The Wishart distribution. In *Multivariate Statistics*, volume 53, pages 302–334. Institute of Mathematical Statistics, 2007.

[Ver18] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.