What is the simplest model problem for the replica trick in random matrix theory?

Jacob A. Zavatone-Veth*

November 10, 2024

Abstract

In these notes, we discuss what may be the simplest random matrix problem to which one can apply the replica trick: the limiting log-determinant of a Wishart matrix. This computation is simple to the point of being unrepresentative of most replica computations, but it introduces some of the core concepts.

Contents

1 Introduction

Consider a d-dimensional Wishart matrix $\mathbf{W}$ with p degrees of freedom. That is, let

$$\mathbf{W} = \frac{1}{p}\mathbf{X}\mathbf{X}^{\top} \tag{1}$$

where $\mathbf{X} \in \mathbb{R}^{d\times p}$ has i.i.d. Gaussian elements $X_{ij} \sim \mathcal{N}(0,1)$. In my previous notes [7], I showed how the replica trick may be used to compute the averages of the minimum and maximum eigenvalues of this matrix in the limit $d,p \to \infty$ with $p/d \to \alpha$. In revisiting these notes following a question from Binxu Wang, it has occurred to me that there is a question one can pose about Wishart matrices with an answer that can be obtained via a closely related but slightly simpler replica computation. This question is extremely simple: for $\alpha > 1$, what is the limiting value of

$$\phi(\alpha) = \lim_{\substack{d,p\to\infty \\ p/d\to\alpha}} \mathbb{E}_{\mathbf{W}} \frac{1}{d} \log\det(\mathbf{W})? \tag{2}$$

*Society of Fellows and Center for Brain Science, Harvard University
jzavatoneveth@fas.harvard.edu

Though this question may seem at first glance somewhat contrived, it is in fact (a) in itself a problem one encounters in the study of high-dimensional Bayesian inference and (b) it leads to a replica computation that is closely related to the computation of the full spectrum of eigenvalues of $\mathbf{W}$, but is somewhat simpler as all integrals are real. For details of the spectral density computation, see [1, 2, 8] or Song Mei's excellent blog posts [3]. Before diving in, I note that I will be somewhat sparing with references and general exposition here; see my longer previous note for more details [7].

In the replica approach to compuuting the log-determinant, we use the Gaussian integral identity

$$\det(\mathbf{W})^{-1/2} = \int_{\mathbb{R}^d} \frac{d\mathbf{u}}{(2\pi)^{d/2}} e^{-\frac{1}{2}\mathbf{u}^\top \mathbf{W}\mathbf{u}} \tag{3}$$

which is well-defined as $\mathbf{W}$ is invertible with probability 1 so long as $p \geq d$. From a statistical mechanics perspective, we can then view the problem as a Gaussian spin model with interaction matrix $\mathbf{W}$, with a Gibbs distribution over states given by

$$p_{\mathbf{W}}(\mathbf{u}) = \frac{1}{(2\pi)^{d/2}\det(\mathbf{W})^{-1/2}} e^{-\frac{1}{2}\mathbf{u}^\top \mathbf{W}\mathbf{u}} \tag{4}$$

for each fixed realization of $\mathbf{W}$ [4]. In this picture, $\det(\mathbf{W})^{-1/2}$ is the 'partition function,' and the quantity of interest

$$\phi(\alpha) = -\lim_{\substack{d,p\to\infty \\ p/d\to\alpha}} \mathbb{E}_{\mathbf{W}} \frac{2}{d} \log \int_{\mathbb{R}^d} \frac{d\mathbf{u}}{(2\pi)^{d/2}} e^{-\frac{1}{2}\mathbf{u}^\top \mathbf{W}\mathbf{u}} \tag{5}$$

is the limiting value of the free energy, averaged over realizations of the interaction matrix $\mathbf{W}$. In this language, $\mathbf{W}$ is referred to as 'quenched' randomness because each realization of $\mathbf{W}$ defines a distribution over the state variable $\mathbf{u}$. This physical perspective can yield useful intuition, though it is not necessary.

2 A warmup: the annealed average

We must now contend with the need to compute the average of a logarithm, which is the challenge by which the replica trick is motivated. Before jumping into the full replica computation, we first consider what is known as an 'annealed' computation, which we will eventually find to be exact. The 'annealed' approximation starts with the observation that, by Jensen's inequality, we have

$$\mathbb{E}_{\mathbf{W}} \log\det(\mathbf{W}) \geq \log \mathbb{E}_{\mathbf{W}} \det(\mathbf{W}) \tag{6}$$

by the concavity of the logarithm. Therefore, we have

$$\phi(\alpha) \leq \phi_{\text{ann}}(\alpha), \tag{7}$$

where

$$\phi_{\text{ann}}(\alpha) = -\lim_{\substack{d,p\to\infty \\ p/d\to\alpha}} \frac{2}{d} \log \mathbb{E}_{\mathbf{W}} \int_{\mathbb{R}^d} \frac{d\mathbf{u}}{(2\pi)^{d/2}} e^{-\frac{1}{2}\mathbf{u}^\top \mathbf{W}\mathbf{u}}. \tag{8}$$

We now evaluate $\phi_{\text{ann}}(\alpha)$, which is straightforward to compute. Interchanging the average over $\mathbf{W}$ with the integral over $\mathbf{u}$ and using the fact that the columns of $\mathbf{X}$ are i.i.d. Gaussian vectors $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$, we can compute

$$\mathbb{E}_{\mathbf{W}} e^{-\frac{1}{2}\mathbf{u}^\top \mathbf{W}\mathbf{u}} = \mathbb{E}_{\mathbf{X}} e^{-\frac{1}{2}\mathbf{u}^\top \mathbf{X}\mathbf{X}^\top \mathbf{u}/p} = \left(1 + \frac{1}{p}\|\mathbf{u}\|^2\right)^{-p/2}. \tag{9}$$

Then, we are left with the integral

$$\int_{\mathbb{R}^d} \frac{d\mathbf{u}}{(2\pi)^{d/2}} \left(1+\frac{1}{p}\|\mathbf{u}\|^2\right)^{-p/2}. \tag{10}$$

To evaluate this integral, we use spherical coordinates with $\|\mathbf{u}\| = u$. The angular integral is trivial as the integrand is radial, and yields the surface area of the $(d-1)$-sphere of unit radius: $2\pi^{d/2}/\Gamma(d/2)$. Thus, we have

$$\frac{2}{2^{d/2}\Gamma(d/2)} \int_0^\infty du\, u^{d-1} \left(1+\frac{1}{p}u^2\right)^{-p/2}. \tag{11}$$

Making the change of variables $q = u^2/d$, this becomes

$$\left(\frac{d}{2}\right)^{d/2} \frac{1}{\Gamma(d/2)} \int_0^\infty \frac{dq}{q} e^{dg(q)} \tag{12}$$

where

$$g(q) = \frac{1}{2}\log q - \frac{1}{2}\alpha \log\left(1+\frac{1}{\alpha}q\right). \tag{13}$$

To summarize our progress, we have

$$\phi_{\text{ann}}(\alpha) = -\lim_{d\to\infty} \frac{2}{d} \log\left[\left(\frac{d}{2}\right)^{d/2} \frac{1}{\Gamma(d/2)}\right] - \lim_{d\to\infty} \frac{2}{d} \log \int_0^\infty \frac{dq}{q} e^{-dg(q)}. \tag{14}$$

The first term is easy to compute using Stirling's approximation for the Γ function,

$$\left(\frac{d}{2}\right)^{d/2} \frac{1}{\Gamma(d/2)} = \frac{1}{2}\left(\frac{d}{\pi}\right)^{1/2} e^{d/2}\left[1+\mathcal{O}\left(\frac{1}{d}\right)\right], \tag{15}$$

which yields

$$\lim_{d\to\infty} \frac{2}{d} \log\left[\left(\frac{d}{2}\right)^{d/2} \frac{1}{\Gamma(d/2)}\right] = 1. \tag{16}$$

Considering the integral over q, the form of the integrand suggests that we may apply Laplace's method. For $q \in [0,\infty)$, and $\alpha > 1$, $g(q)$ is a non-positive function with $g(q) \to -\infty$ as $q \downarrow 0$ or $q \to \infty$. It has a unique maximum at

$$q_* = \frac{\alpha}{\alpha-1}, \tag{17}$$

where it takes the value

$$g(q_*) = -\frac{1}{2}(\alpha-1)\log\frac{\alpha}{\alpha-1}. \tag{18}$$

This can be seen from the fact that q_* is the only positive real solution to $0 = g'(q_*) = \frac{1}{2q} - \frac{1}{2(1+q/\alpha)}$ and the fact that

$$g''(q_*) = -\frac{1}{2}\left(\frac{\alpha-1}{\alpha}\right)^3. \tag{19}$$

Thus, from Laplace's method, we have that

$$\lim_{d\to\infty} \frac{2}{d} \log \int_0^\infty \frac{dq}{q} e^{-dg(q)} = 2g(q_*). \tag{20}$$

Collecting these results, we have that

$$\phi_{\text{ann}}(\alpha) = (\alpha-1)\log\frac{\alpha}{\alpha-1} - 1. \tag{21}$$

3 The replica trick

We claim that the annealed computation in this case gives not only a lower bound on ϕ but in fact precisely determines it. How could we show this? More generally, how could we handle the average of the logarithm in cases where the annealed average is not exact? The replica trick provides a physicist's answer to these questions.

Fundamentally, the replica trick is based on the variants of identity

$$\mathbb{E}\log Z = \lim_{n\to 0}\frac{1}{n}\log \mathbb{E}Z^n, \tag{22}$$

for suitably well-behaved positive random variables Z. From this starting point, the replica trick makes two important assumptions:

1. It suffices to compute integer moments $\mathbb{E}Z^n$ for $n \in \mathbb{N}$, and then analytically continue to $n \to 0$.
2. The 'thermodynamic' limit $d \to \infty$ can be interchanged with the limit $n \to 0$.

These two assumptions are not always rigorously justified. The first is particularly subtle, as there are important problems for which the analytic continuation of $\mathbb{E}Z^n$ is not unique. For some references, see my other notes [7]. For the simple problem at hand, however, we will not encounter any such complications.

With this in mind, we compute

$$\phi_n(\alpha) = -\lim_{\substack{d,p\to\infty \\ p/d\to\alpha}} \frac{2}{d}\log \mathbb{E}_{\mathbf{W}}\det(\mathbf{W})^{-n/2} \tag{23}$$

for $n \in \mathbb{N}$, with the hope that

$$\phi(\alpha) = \lim_{n\to 0}\frac{\phi_n(\alpha)}{n}. \tag{24}$$

The expectation over $\mathbf{W}$ is easy to evaluate:

$$\mathbb{E}_{\mathbf{W}}\det(\mathbf{W})^{-n/2} = \mathbb{E}_{\mathbf{W}}\left(\int_{\mathbb{R}^d}\frac{d\mathbf{u}}{(2\pi)^{d/2}}e^{-\frac{1}{2}\mathbf{u}^\top\mathbf{W}\mathbf{u}}\right)^n \tag{25}$$

$$= \mathbb{E}_{\mathbf{W}}\int\prod_{a=1}^n\frac{d\mathbf{u}_a}{(2\pi)^{d/2}}e^{-\frac{1}{2}\sum_{a=1}^n\mathbf{u}_a^\top\mathbf{W}\mathbf{u}_a} \tag{26}$$

$$= \int\prod_{a=1}^n\frac{d\mathbf{u}_a}{(2\pi)^{d/2}}\mathbb{E}_{\mathbf{X}}e^{-\frac{1}{2p}\sum_{a=1}^n\mathbf{u}_a^\top\mathbf{X}\mathbf{X}^\top\mathbf{u}_a} \tag{27}$$

$$= \int\prod_{a=1}^n\frac{d\mathbf{u}_a}{(2\pi)^{d/2}}\left(\mathbb{E}_{\mathbf{x}\sim\mathcal{N}(\mathbf{0},\mathbf{I}_d)}e^{-\frac{1}{2p}\sum_{a=1}^n\mathbf{u}_a^\top\mathbf{x}\mathbf{x}^\top\mathbf{u}_a}\right)^p \tag{28}$$

$$= \int\prod_{a=1}^n\frac{d\mathbf{u}_a}{(2\pi)^{d/2}}\det\left(\mathbf{I}_d + \frac{1}{p}\sum_{a=1}^n\mathbf{u}_a\mathbf{u}_a^\top\right)^{-p/2}. \tag{29}$$

Using the Weinstein–Aronszajn identity, we can re-write the $d\times d$ determinant as a $n\times n$ determinant, whose size remains constant as $d\to\infty$:

$$\det\left(\mathbf{I}_d + \frac{1}{p}\sum_{a=1}^n\mathbf{u}_a\mathbf{u}_a^\top\right) = \det\left(\mathbf{I}_n + \frac{1}{\alpha}\mathbf{Q}\right), \tag{30}$$

where we have defined the Gram matrix

$$Q_{ab} = \frac{1}{d}\mathbf{u}_a^\top \mathbf{u}_b. \tag{31}$$

Now, in analogy to what we did in the annealed computation, we would like to change variables from $\mathbf{u}_a$ to Q_{ab}. To do this, we use a standard trick: we write

$$1 = \int d\mathbf{Q} \prod_{a\leq b} \delta\left(Q_{ab} - \frac{1}{d}\mathbf{u}_a^\top \mathbf{u}_b\right), \tag{32}$$

where the integral is taken over all positive-definite $n \times n$ matrices. Then, we have

$$\mathbb{E}_{\mathbf{W}} \det(\mathbf{W})^{-n/2} = \int d\mathbf{Q} \det\left(\mathbf{I}_n + \frac{1}{\alpha}\mathbf{Q}\right)^{-p/2} \int \prod_{a=1}^{n} \frac{d\mathbf{u}_a}{(2\pi)^{d/2}} \prod_{a\leq b} \delta\left(Q_{ab} - \frac{1}{d}\mathbf{u}_a^\top \mathbf{u}_b\right). \tag{33}$$

The integral

$$S(\mathbf{Q}) = \int \prod_{a=1}^{n} \frac{d\mathbf{u}_a}{(2\pi)^{d/2}} \prod_{a\leq b} \delta\left(Q_{ab} - \frac{1}{d}\mathbf{u}_a^\top \mathbf{u}_b\right) \tag{34}$$

is precisely the normalized volume of states with a fixed overlap $\mathbf{Q}$, and thus represents an 'entropic' contribution. Put another way, we want to know the Lebesgue measure of the set of n vectors $\mathbf{u}_a \in \mathbb{R}^d$ with Gram matrix $\mathbf{Q}$. There are a variety of approaches by which this integral could be evaluated asymptotically for large d, the most commonly-used one in replica computations being to introduce further auxiliary variables (see [7]). In this case, however, we may compute the entropic term directly; in Appendix A we show that it evaluates to

$$\int \prod_{a=1}^{n} \frac{d\mathbf{u}_a}{(2\pi)^{d/2}} \prod_{a\leq b} \delta\left(Q_{ab} - \frac{1}{d}\mathbf{u}_a^\top \mathbf{u}_b\right) = C_{d,n} \det(\mathbf{Q})^{(d-n-1)/2} \tag{35}$$

for

$$C_{d,n} = \left(\frac{d}{2\pi}\right)^{dn/2} \prod_{a=1}^{n} \frac{\pi^{(d-a+1)/2}}{\Gamma[(d-a+1)/2]}, \tag{36}$$

which reduces to the corresponding result in the annealed case when we set $n = 1$. Thus, we have

$$\mathbb{E}_{\mathbf{W}} \det(\mathbf{W})^{-n/2} = C_{d,n} \int d\mathbf{Q}\, e^{dg(\mathbf{Q})} \det(\mathbf{Q})^{-(n+1)/2} \tag{37}$$

where

$$g(\mathbf{Q}) = \frac{1}{2}\log\det(\mathbf{Q}) - \frac{1}{2}\alpha \log\det\left(\mathbf{I}_n + \frac{1}{\alpha}\mathbf{Q}\right). \tag{38}$$

Then, we have

$$\lim_{\substack{d,p\to\infty \\ p/d\to\alpha}} \frac{1}{d}\log\det(\mathbf{W})^{-n/2} = \lim_{d\to\infty} \frac{1}{d}\log C_{d,n} + \lim_{d\to\infty}\frac{1}{d}\log \int d\mathbf{Q}\, e^{dg(\mathbf{Q})} \det(\mathbf{Q})^{-(n+1)/2}. \tag{39}$$

Based on our discussion in the annealed case, it is easy to see that for any fixed n we have

$$\lim_{d\to\infty} \frac{1}{d} \log C_{d,n} = \frac{n}{2}. \tag{40}$$

Considering the integral term

$$\lim_{d\to\infty} \frac{1}{d} \log \int d\mathbf{Q}\, e^{dg(\mathbf{Q})} \det(\mathbf{Q})^{-n}, \tag{41}$$

we observe that

$$\frac{\delta g}{\delta \mathbf{Q}} = \frac{1}{2}\mathbf{Q}^{-1} - \frac{1}{2}\left(\mathbf{I}_n + \frac{1}{\alpha}\mathbf{Q}\right)^{-1} \tag{42}$$

from which we can see that $g(\mathbf{Q})$ has a single stationary point among the symmetric positive-definite matrices at

$$\mathbf{Q}_* = \frac{\alpha}{\alpha-1}\mathbf{I}_n. \tag{43}$$

We claim that for any integer n this is a maximum. To do so, we consider the Hessian

$$\frac{\partial^2 g}{\partial Q_{ab}\partial Q_{cd}} = \frac{\partial}{\partial Q_{cd}}\left[\frac{1}{2}\mathbf{Q}^{-1} - \frac{1}{2}\left(\mathbf{I}_n + \frac{1}{\alpha}\mathbf{Q}\right)^{-1}\right]_{ab} \tag{44}$$

$$= -\frac{1}{2}(\mathbf{Q}^{-1})_{ac}(\mathbf{Q}^{-1})_{bd} + \frac{1}{2\alpha}\left(\mathbf{I}_n + \frac{1}{\alpha}\mathbf{Q}\right)^{-1}_{ac}\left(\mathbf{I}_n + \frac{1}{\alpha}\mathbf{Q}\right)^{-1}_{bd}, \tag{45}$$

which when evaluated at $\mathbf{Q}_*$ gives

$$\left.\frac{\partial^2 g}{\partial Q_{ab}\partial Q_{cd}}\right|_{\mathbf{Q}=\mathbf{Q}_*} = -\frac{1}{2}\left(\frac{\alpha-1}{\alpha}\right)^3 \delta_{ac}\delta_{bd}. \tag{46}$$

Thus, if one vectorizes the matrix $\mathbf{Q}$, the Hessian at the stationary point is diagonal and has negative diagonal entries. Then, as in the annealed case, the integral is dominated by this stationary point—in the present setting this is justified by the method of steepest descent, a wide-ranging generalization of Laplace's method—and we have

$$\lim_{d\to\infty} \frac{1}{d} \log \int d\mathbf{Q}\, e^{dg(\mathbf{Q})} \det(\mathbf{Q})^{-(n+1)/2} = g(\mathbf{Q}_*) \tag{47}$$

$$= -\frac{n}{2}(\alpha-1)\log\frac{\alpha}{\alpha-1}. \tag{48}$$

Combining these results, we find that

$$\phi_n(\alpha) = n\phi_{\text{ann}}(\alpha), \tag{49}$$

from which we can see that

$$\phi(\alpha) = \lim_{n\to 0}\frac{1}{n}\phi_n(\alpha) = \lim_{n\to 0}\phi_{\text{ann}}(\alpha) = \phi_{\text{ann}}(\alpha). \tag{50}$$

In this case, the analytic continuation to $n \to 0$ is clearly unique; the formal justification for this comes from Carlson's theorem. This justifies our claim that the annealed result is exact.

From a physical perspective, what we have found is that the replicas become uncorrelated in the limit $d \to \infty$. At any fixed n, this can be seen more directly from the interpretation of the Gaussian integral as a system of interacting units.[1] Let $\langle \cdot \rangle$ denote averaging with respect to the Gibbs measure for a fixed realization of $\mathbf{W}$, *i.e.*,

$$\langle \cdot \rangle = \int d\mathbf{u}\, p_{\mathbf{W}}(\mathbf{u})(\cdot). \tag{51}$$

Then, if we define $\langle \cdot \rangle_n$ to be the average over the replicated system for a fixed $n \in \mathbb{N}$,

$$\langle \cdot \rangle_n = \int \prod_{a=1}^{n} d\mathbf{u}_a \prod_{a=1}^{n} p_{\mathbf{W}}(\mathbf{u})(\cdot), \tag{52}$$

we can see that for any distinct replicas $b \neq a$ we have

$$\left\langle \frac{1}{d}\mathbf{u}_a^\top \mathbf{u}_b \right\rangle = \frac{1}{d}\langle \mathbf{u}_a \rangle^\top \langle \mathbf{u}_b \rangle \tag{53}$$

$$= \frac{1}{d}\|\langle \mathbf{u} \rangle\|^2 \tag{54}$$

as for a fixed $\mathbf{W}$ the replicas are uncoupled and equivalent. But, observing that the Gibbs distribution is symmetric in the sense that

$$p_{\mathbf{W}}(-\mathbf{u}) = p_{\mathbf{W}}(\mathbf{u}) \tag{55}$$

for any $\mathbf{u}$, we clearly have

$$\langle \mathbf{u} \rangle = \mathbf{0} \tag{56}$$

for any $\mathbf{W}$. Now, in the limit $d \to \infty$, the argument above implies that the saddle-point value gives the average of $\mathbf{Q}$ with respect to the Gibbs distribution, *i.e.*,

$$(Q_*)_{ab} = \lim_{d \to \infty} \mathbb{E}_{\mathbf{W}} \left\langle \frac{1}{d}\mathbf{u}_a^\top \mathbf{u}_b \right\rangle. \tag{57}$$

But, we have just seen that the right-hand-side of this expression vanishes for $b \neq a$, which is consistent with our previous finding that $\mathbf{Q}_*$ is diagonal.

4 Concluding thoughts

To conclude, we have shown that

$$\phi(\alpha) = (\alpha - 1)\log\frac{\alpha}{\alpha - 1} - 1. \tag{58}$$

We note that this result may also be obtained using the Marchenko-Pastur theorem, which gives the asymptotic distribution of eigenvalues of a Wishart matrix. This approach is detailed in a StackExchange post,[2] which, interestingly enough, uses the replica trick to handle the logarithm!

[1] Two comments are in order: First, I learned this argument from Haim Sompolinsky in a lecture on [5]. Second, I will not worry too much about limits in n here.

[2] https://math.stackexchange.com/questions/2250884/integration-over-the-marchenko-pastur-distribution

Throughout this note, I have tried to be as explicit as possible, and spell out all intermediate steps. This contrasts with most expositions of the replica method—including my own previous notes [7]—which sweep some things under the rug. However, thanks to the simplicity of this example, we have not encountered many of the challenges present in most replica computations. In particular, the saddle point in $\mathbf{Q}$ could be computed exactly at any n without making an *Ansatz* on the form of the matrix. This is not representative of most replica computations. For a slightly more complicated example, see my existing notes on the minimum and maximum eigenvalues of Wishart matrices [7].

Acknowledgements

I again thank Binxu Wang for posing the question about my older notes that inspired this example.

A Evaluation of the entropic integral

In this appendix, we show that for an $n \times n$ real symmetric positive-definite matrix $\mathbf{Q}$ one has

$$I_{d,n}(\mathbf{Q}) = \int \prod_{a=1}^{n} d\mathbf{u}_a \prod_{a\leq b} \delta(Q_{ab} - \mathbf{u}_a^\top \mathbf{u}_b) = \left[\prod_{a=1}^{n} \frac{\pi^{(d-a+1)/2}}{\Gamma[(d-a+1)/2]} \right] \det(\mathbf{Q})^{(d-n-1)/2} \tag{59}$$

for $\mathbf{u}_a \in \mathbb{R}^d$, which leads to the expression for the entropic integral given in the main text after a re-scaling. This is closely related to the derivation of the probability density function of a Wishart distribution, except for the fact that integration over $\mathbf{u}_a$ is performed with respect to a flat measure rather than a Gaussian one.

With this relationship in mind, we will give a proof that closely follows Wishart's derivation of his eponymous probability density [6]. An analogous but slightly algebraically simpler proof follows by noting that we may work in coordinates where $\mathbf{Q}$ is diagonal. Putting that idea aside, we proceed by induction on n. We begin with the base case $n = 1$:

$$I_{d,n=1}(Q) = \int_{\mathbb{R}^d} d\mathbf{u}_a \, \delta(Q - \|\mathbf{u}\|^2). \tag{60}$$

This integral of course gives the surface area of the $(d-1)$-sphere of radius $Q^{1/2}$, though we must be careful to account for the fact that the constraint is on the square of the radius $r = \|\mathbf{u}\|$: $\delta(Q-r^2)$, which introduces a Jacobian factor of $1/|\partial_r(Q-r^2)| = 1/(2r)$. Thus, using the well-known fact that the surface area of the unit sphere in d dimensions is $2\pi^{d/2}/\Gamma(d/2)$, we have

$$I_{d,n=1}(Q) = \frac{\pi^{d/2}}{\Gamma(d/2)} Q^{(d-2)/2}, \tag{61}$$

which is the desired result.

We now consider some $n > 1$. As it will prove useful, we introduce the notation $\mathbf{Q}_m$ to represent the $m \times m$ matrix with elements Q_{ab} for $a, b = 1, \dots, m$, which by assumption is symmetric and positive-definite. On the induction hypothesis, we have that

$$\frac{I_{d,n}(\mathbf{Q}_n)}{I_{d,n-1}(\mathbf{Q}_{n-1})} = \frac{\pi^{(d-n+1)/2}}{\Gamma[(d-n+1)/2]} \frac{\det(\mathbf{Q}_n)^{(d-n-1)/2}}{\det(\mathbf{Q}_{n-1})^{(d-n)/2}}. \tag{62}$$

To show that this recurrence holds, we proceed by isolating the integral over $\mathbf{u}_n$:

$$I_{d,n}(\mathbf{Q}_n) = \int \prod_{a=1}^{n-1} d\mathbf{u}_a \prod_{1\leq a\leq b\leq n-1} \delta(Q_{ab} - \mathbf{u}_a^\top \mathbf{u}_b) \int d\mathbf{u}_n \prod_{a=1}^{n} \delta(Q_{an} - \mathbf{u}_a^\top \mathbf{u}_n). \tag{63}$$

Now we can consider the integral

$$\int d\mathbf{u}_n \prod_{a=1}^{n} \delta(Q_{an} - \mathbf{u}_a^\top \mathbf{u}_n) \tag{64}$$

in isolation, with $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ fixed vectors satisfying $\mathbf{u}_a \cdot \mathbf{u}_b = Q_{ab}$ for $a, b = 1, \dots, n-1$. As we assume $\mathbf{Q}_{n-1}$ is full-rank, the set of vectors $\mathbf{u}_a$ for $a = 1, \dots, n-1$ are linearly independent, so we may decompose $\mathbf{u}_n$ into a linear combination of those vectors plus a component $\mathbf{u}_n^\perp$ orthogonal to their span:

$$\mathbf{u}_n = \sum_{a=1}^{n-1} v_a \mathbf{u}_a + \mathbf{u}_n^\perp. \tag{65}$$

We would like to change variables in the integral to this representation. As the orthogonal complement is $d-(n-1)$-dimensional, we may write $\mathbf{u}_n^\perp = \mathbf{P}_{n-1}\mathbf{z}_n$, where $\mathbf{z}_n \in \mathbb{R}^{d-n+1}$ and $\mathbf{P}_{n-1} \in \mathbb{R}^{d,d-n+1}$ is a projector satisfying $\mathbf{u}_a^\top \mathbf{P}_{n-1} = \mathbf{0}$ for all $a = 1,\ldots,n-1$ and $\mathbf{P}_{n-1}^\top \mathbf{P}_{n-1} = \mathbf{I}_{d-n+1}$. Then, we have the Jacobian

$$\frac{\partial \mathbf{u}_n}{\partial(\mathbf{v},\mathbf{z})} = \left(\mathbf{u}_1,\cdots,\mathbf{u}_{n-1},\mathbf{P}_{n-1}\right). \tag{66}$$

Using the orthogonality constraint, we can compute that

$$\left|\det \frac{\partial \mathbf{u}_n}{\partial(\mathbf{v},\mathbf{z})}\right| = \det(\mathbf{Q}_{n-1})^{1/2}. \tag{67}$$

With this decomposition, we have

$$\|\mathbf{u}_n\|^2 = \sum_{a,b=1}^{n-1} Q_{ab} v_a v_b + \|\mathbf{u}_n^\perp\|^2 \tag{68}$$

and

$$\mathbf{u}_n \cdot \mathbf{u}_a = \sum_{b=1}^{n-1} Q_{ab} v_b. \tag{69}$$

Thus, the integral over $\mathbf{u}_n$ becomes

$$\det(\mathbf{Q}_{n-1})^{1/2} \int_{\mathbb{R}^{n-1}} d\mathbf{v} \int_{\mathbb{R}^{d-n+1}} d\mathbf{z}\, \delta\left(Q_{nn} - \mathbf{v}^\top \mathbf{Q}_{n-1}\mathbf{v} - \|\mathbf{z}\|^2\right) \prod_{a=1}^{n-1} \delta\left(Q_{an} - \sum_{b=1}^{n-1} Q_{ab} v_b\right) \tag{70}$$

We now make a further change of variables

$$\mathbf{v} = \mathbf{Q}_{n-1}^{-1}\mathbf{s} \tag{71}$$

which has Jacobian determinant $\det(\mathbf{Q}_{n-1})^{-1}$, yielding

$$\det(\mathbf{Q}_{n-1})^{-1/2} \int_{\mathbb{R}^{n-1}} d\mathbf{s} \int_{\mathbb{R}^{d-n+1}} d\mathbf{z}\, \delta\left(Q_{nn} - \mathbf{s}^\top \mathbf{Q}_{n-1}^{-1}\mathbf{s} - \|\mathbf{z}\|^2\right) \prod_{a=1}^{n-1} \delta\left(Q_{an} - s_a\right). \tag{72}$$

The integral over $\mathbf{s}$ is now of course easy to evaluate, giving

$$\det(\mathbf{Q}_{n-1})^{-1/2} \int_{\mathbb{R}^{d-n+1}} d\mathbf{z}\, \delta\left(Q_{nn} - \sum_{a,b=1}^{n-1} (Q_{n-1}^{-1})_{ab} Q_{an} Q_{bn} - \|\mathbf{z}\|^2\right). \tag{73}$$

The Schur complement formula implies that

$$Q_{nn} - \sum_{a,b=1}^{n-1} (Q_{n-1}^{-1})_{ab} Q_{an} Q_{bn} = \frac{\det(\mathbf{Q}_n)}{\det(\mathbf{Q}_{n-1})} \tag{74}$$

leaving us with

$$\det(\mathbf{Q}_{n-1})^{-1/2} \int_{\mathbb{R}^{d-n+1}} d\mathbf{z}\, \delta\left(\frac{\det(\mathbf{Q}_n)}{\det(\mathbf{Q}_{n-1})} - \|\mathbf{z}\|^2\right). \tag{75}$$

This is, at last, another integral over the surface of a sphere, so, accounting for the fact that the constraint is on the square of the radius as we did with $n = 1$, we have

$$\det(\mathbf{Q}_{n-1})^{-1/2}\frac{\pi^{(d-n+1)/2}}{\Gamma[(d-n+1)/2]}\left(\frac{\det(\mathbf{Q}_n)}{\det(\mathbf{Q}_{n-1})}\right)^{(d-n-1)/2} = \frac{\pi^{(d-n+1)/2}}{\Gamma[(d-n+1)/2]}\frac{\det(\mathbf{Q}_n)^{(d-n-1)/2}}{\det(\mathbf{Q}_{n-1})^{(d-n)/2}}. \tag{76}$$

Therefore, we find that

$$I_{d,n}(\mathbf{Q}_n) = \int \prod_{a=1}^{n-1} d\mathbf{u}_a \prod_{1\leq a\leq b\leq n-1} \delta(Q_{ab} - \mathbf{u}_a^\top \mathbf{u}_b)\frac{\pi^{(d-n+1)/2}}{\Gamma[(d-n+1)/2]}\frac{\det(\mathbf{Q}_n)^{(d-n-1)/2}}{\det(\mathbf{Q}_{n-1})^{(d-n)/2}} \tag{77}$$

$$= \frac{\pi^{(d-n+1)/2}}{\Gamma[(d-n+1)/2]}\frac{\det(\mathbf{Q}_n)^{(d-n-1)/2}}{\det(\mathbf{Q}_{n-1})^{(d-n)/2}} I_{d,n-1}(\mathbf{Q}_{n-1}), \tag{78}$$

which proves the claimed recurrence, and thus concludes the proof.

References

[1] M. Advani, S. Lahiri, and S. Ganguli, “Statistical mechanics of complex neural systems and high dimensional data”, Journal of Statistical Mechanics: Theory and Experiment **2013**, P03014 (2013).

[2] S. F. Edwards and R. C. Jones, “The eigenvalue spectrum of a large symmetric random matrix”, Journal of Physics A: Mathematical and General **9**, 1595–1603 (1976), https://doi.org/10.1088/0305-4470/9/10/011.

[3] S. Mei, *Replica method and random matrices*, online blog post, URL: https://meisong541.github.io/jekyll/update/2019/08/04/Replica_method_1.html and https://meisong541.github.io/jekyll/update/2019/09/12/Replica_method_2.html, 2019.

[4] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond: an introduction to the replica method and its applications* (World Scientific Publishing Company, 1987).

[5] H. J. Sommers, A. Crisanti, H. Sompolinsky, and Y. Stein, “Spectrum of large random asymmetric matrices”, Phys. Rev. Lett. **60**, 1895–1898 (1988), https://link.aps.org/doi/10.1103/PhysRevLett.60.1895.

[6] J. Wishart, “The generalised product moment distribution in samples from a normal multivariate population”, Biometrika **20A**, 32–52 (1928), https://doi.org/10.1093/biomet/20A.1-2.32.

[7] J. A. Zavatone-Veth, *Lecture notes on the replica method for Wishart matrix eigenvalues*, Lecture notes, available online, 2022, https://jzv.io/assets/pdf/wishart_eigenvalue_and_gaussian_data_pca_replica_notes.pdf.

[8] J. A. Zavatone-Veth and C. Pehlevan, “Replica method for eigenvalues of real Wishart product matrices”, arXiv, 10.48550/ARXIV.2209.10499 (2022), https://arxiv.org/abs/2209.10499.