The Kac-Rice formula : basic definitions and application

Antoine Maillard

February 12, 2019

1 The Kac-Rice formula

1.1 The area formula

The Kac-Rice formula is not derived from any involved probabilistic tools, but rather is a consequence of a deeper geometric result, called the area formula (itself a consequence of the coreao formula), described in more generality by Federer [Fed59], and stated for instance in [AW09]. This formula is the generalization of the following non-rigorous intuition: for a smooth function $f : \mathbb{R} \to \mathbb{R}$, and $T \subset \mathbb{R}$, denoting $N_f(u,T)$ the number of solutions to the equation $f(t) = u$ with $t \in T$, one would want to write informally:

$$N_f(u,T) = \int_{f(T)} \delta\left(v-u\right) \mathrm{d}v, \tag{1}$$

$$= \int_T \delta\left(f(t)-u\right) |f'(t)| \mathrm{d}t. \tag{2}$$

The area formula generalizes and makes rigorous this intuition, by showing a weak version of this equality. We follow here the statement of [AW09].

Proposition 1.1 (Area formula) *Let $f : U \to \mathbb{R}^d$ be a $\mathcal{C}^1$ function defined on an open subset U of $\mathbb{R}^d$. Assume that the sets of critical values of f has zero Lebesgue measure, and denote $N_f(u,T)$ the number of solutions to the equation $f(t) = u$ with $t \in T$. Then, for any Borel set $T \subset \mathbb{R}^d$, and any $g : \mathbb{R}^d \to \mathbb{R}$ continuous and bounded:*

$$\int_{\mathbb{R}^d} g(u) N_f(u,T) = \int_T |\det f'(t)| \; g(f(t)) \,\mathrm{d}t. \tag{3}$$

1.2 Informal derivation of Kac-Rice

Consider a compact manifold $\mathcal{M}$ of dimension n, and a random function $f : \mathcal{M} \to \mathbb{R}$. We want to use the area formula 1.1 to estimate the moments of the number of critical points of f. Given the hypotheses of Prop. 1.1, a reasonable hypothesis is to assume that f is almost surely a *Morse* function, i.e. that all its critical points are non-degenerate. Since $\mathcal{M}$ is compact, one easily deduces that the number of critical points of f is finite. [1] For any $k \in \mathbb{N}$ and Borel set $B \subseteq \mathbb{R}$, we define $\mathrm{Crit}_{f,k}(B)$ to be the number of critical points $x \in \mathcal{M}$ of f such that $f(x) \in B$ and such that the index of $\mathrm{Hess}\, f(x)$ is equal to k. Informally, one can write:

$$\mathrm{Crit}_{f,k}(B) = \int_{\mathcal{M}} \mu_{\mathcal{M}}(\mathrm{d}x) \delta\left(\mathrm{grad}\, f(x)\right) |\det \mathrm{Hess}\, f(x)| \; \mathbb{1}\left[f(x) \in B, \, \mathrm{i}\left(\mathrm{Hess}\, f(x)\right) = k\right] \tag{4}$$

Taking the expectation of this equality, one obtains the Kac-Rice formula:

[1] Note that the numbers of critical points of f of different indices are related to the topology of $\mathcal{M}$ by the Morse inequalities, see [MSWW63] for a review on Morse theory.

Proposition 1.2 (Kac-Rice formula) *Denote* $\varphi_{\mathrm{grad}\, f(x)}(0)$ *the density of* $\mathrm{grad}\, f(x)$ *taken at* 0. *Then:*

$$\mathbb{E}\,\mathrm{Crit}_{f,k}(B) = \int_{\mathcal{M}} \mu_{\mathcal{M}}(\mathrm{d}x)\mathbb{E}\left[|\det\, \mathrm{Hess}\, f(x)|\ \mathbb{1}\left[f(x)\in B,\, \mathrm{i}\left(\mathrm{Hess}\, f(x)\right)=k\right]|\mathrm{grad}\, f(x)=0\right]\varphi_{\mathrm{grad}\, f(x)}(0).$$

Remarks :

1. The rigorous derivation of this formula is much more involved, as one has to start from the weak equality of Prop. 1.1 and use continuity arguments in order to obtain an equality exactly at $u=0$, see [AW09]. This is one reason why the formula is mainly stated for Gaussian random fields, for which these arguments are easier to prove. The second is that in general, conditional expectations of non-Gaussian random variables are intractable. Under many technical conditions, one can however derive non-Gaussian versions of the Kac-Rice formula, stated for example as Thm.12.1.1 of [AT09] or Thm.6.7 of [AW09].

2. The difficulty of the Kac-Rice formula comes from the model of random matrices that arise from the distribution of the Hessian conditioned by the gradient being 0. Even for Gaussian random fields, this is in general a heavily correlated Gaussian random matrix, for which very few results exist.

3. The Kac-Rice formula can be stated to compute higher moments of the variable $\mathrm{Crit}_{f,k}(B)$ as well (Thm 6.3 of [AW09]), and can therefore be used to compute the second moment of the complexity (see [S^+17] for an application to the spherical p-spin model) as well as perform heuristic replica calculations for the quenched complexity (see for instance [RABC19]).

2 An application: annealed complexity of the pure spherical p-spin model

We essentially detail here a calculation performed by physicists, and made rigorous in [AAČ13]. We follow here the derivation of this last paper. For anterior theoretical physics derivation of the complexity of similar models using the Kac-Rice formula, one can for instance read the works of Fyodorov: [Fyo04], [FW07].

2.1 Statement of the problem

Consider $N\geq 1$, $p\geq 3$, and define the function $f_{N,p}$ on the unit sphere $\mathbb{S}^{N-1}$:

$$f_{N,p}(\sigma) \triangleq \sum_{1\leq i_1,\cdots,i_p\leq N} J_{i_1,\cdots i_p}\sigma_{i_1}\cdots\sigma_{i_p}, \tag{5}$$

in which $J_{i_1,\cdots i_p} \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$. In physics terms, if $H_{N,p}$ is the Hamiltonian of the spherical p-spin model, one has $f_{N,p}(\sigma)=\frac{1}{\sqrt{N}}H_{N,p}(\sigma)$. For any Borel set $B\subseteq\mathbb{R}$, we want to compute the large N limit of the expectation of $\mathrm{Crit}^0_{N,p}(B)$, the number of local minima σ of $f_{N,p}$ such that $f_{N,p}(\sigma)\in\sqrt{N}B$. A direct application of the Kac-Rice formula yields:

$$\mathbb{E}\,\mathrm{Crit}^0_{N,p}(B) = \int_{\mathbb{S}^{N-1}} \mu_N(\mathrm{d}\sigma)\varphi_{\mathrm{grad}\, f_{N,p}(\sigma)}(0)\,\mathbb{E}\left[|\det\, \mathrm{Hess}\, f_{N,p}(\sigma)|\ \mathbb{1}\left[f_{n,p}(\sigma)\in B,\, \mathrm{Hess}\, f_{N,p}(\sigma)\geq 0\right]|\mathrm{grad}\, f_{N,p}(\sigma)=0\right],$$

in which μ_N is the usual surface measure on $\mathbb{S}^{N-1}$. Note that here grad and Hess denote the *Riemannian* gradient and Hessian on the sphere. We denote ∇, ∇^2 the Euclidian gradient and Hessian.

2.2 The distribution of $(f(\sigma), \operatorname{grad} f(\sigma), \operatorname{Hess} f(\sigma))$

We fix $\sigma \in \mathbb{S}^{N-1}$. It is trivial to see that all three variables $(f(\sigma), \operatorname{grad} f(\sigma), \operatorname{Hess} f(\sigma))$ are Gaussian centered random variables. We thus simply compute their correlations. We identify the tangent space $\mathcal{T}_\sigma(\mathbb{S}^{N-1})$ with $\mathbb{R}^{N-1}$. If we denote $P_\sigma^\perp$ the orthogonal projector on $\{\sigma\}^\perp$, one has:

$$\operatorname{grad} f(\sigma) = P_\sigma^\perp \nabla f(\sigma), \tag{6}$$

$$\operatorname{Hess} f(\sigma) = P_\sigma^\perp \nabla^2 f(\sigma) P_\sigma^\perp - \langle \sigma, \nabla f(\sigma)\rangle P_\sigma^\perp. \tag{7}$$

So for instance:

$$\begin{aligned}\mathbb{E}\left[\operatorname{grad} f(\sigma) \operatorname{grad} f(\sigma)^\intercal\right] &= P_\sigma^\perp \mathbb{E}\left[\nabla f(\sigma) \nabla f(\sigma)^\intercal\right] P_\sigma^\perp, \\ &= p P_\sigma^\perp.\end{aligned}$$

Uisng the same kind of calculation, one easily obtains that:

Lemma 2.1 *The joint law of $(f(\sigma), \operatorname{grad} f(\sigma), \operatorname{Hess} f(\sigma))$ is the following:*

$$\begin{cases} f(\sigma) & = Z \\ \operatorname{grad} f(\sigma) & = \sqrt{p}\boldsymbol{g} \\ \operatorname{Hess} f(\sigma) & = \sqrt{2(N-1)p(p-1)} M_{N-1} - pZ \operatorname{Id}_{N-1} \end{cases} \tag{8}$$

in which $Z \sim \mathcal{N}(0,1)$, $\boldsymbol{g} \sim \mathcal{N}(0, \operatorname{Id}_{N-1})$, and M_{N-1} is a GOE matrix of size $(N-1)$ with the convention $\mathbb{E} M_{ij}^2 = \frac{1+\delta_{ij}}{2(N-1)}$. The variables $(Z, \boldsymbol{g}, M_{N-1})$ are pairwise independent.

One can make the following remarks:

1. The distribution of all variables is independent of σ.
2. The variables $(f, \operatorname{Hess} f)$ are independent from $\operatorname{grad} f$, so the conditioning in the Kac-Rice formula will be trivial.
3. From this distribution, one easily obtains:

$$\varphi_{\operatorname{grad} f_{N,p}(\sigma)}(0) = e^{-\frac{N-1}{2}\ln(2\pi p)}.$$

Noting that the volume of the unit spher $V(\mathbb{S}^{N-1}) = 2\pi^{N/2}/\Gamma(N/2)$, one deduces from the Kac-Rice formula:

$$\mathbb{E}\operatorname{Crit}^0_{N,p}(B) = \frac{2\pi^{N/2}}{\Gamma(N/2)} e^{-\frac{N-1}{2}\ln(2\pi p)} \left[2(N-1)p(p-1)\right]^{\frac{N-1}{2}} \mathbb{E}\left[|\det H_{N-1}| \mathbb{1}(H_{N-1} \geq 0,\, z \in \sqrt{N}B)\right], \tag{9}$$

in which $H_{N-1} \triangleq M_{N-1} - \sqrt{\frac{p}{2(N-1)(p-1)}} z$. z is a standard Gaussian variable and M_{N-1} is a GOE matrix of size $N-1$ (see Lemma. 2.1). It is now completely clear that we reduced a differential geometry problem (counting the number of critical points of a random function) to a random matrix theory problem.

2.3 Simplification of the problem

It is clear from Eq. 9 that the following lemma will be useful:

Lemma 2.2 *Let $G \subseteq \mathbb{R}$ a Borel set, $X \sim \mathcal{N}(0,t^2)$ and $M_{N-1} \sim \mathrm{GOE}(N-1)$. Then:*

$$\begin{aligned}&\mathbb{E}\left[|\det(M_{N-1} - X\mathrm{Id}_{N-1})|\mathbb{1}((M_{N-1} - X\mathrm{Id}_{N-1}) \geq 0, X \in G)\right] \\ &\quad = \frac{\Gamma\left(\frac{N}{2}\right)(N-1)^{-\frac{N}{2}}}{\sqrt{\pi t^2}} \mathbb{E}_{\mathrm{GOE}(N)}\left[e^{\frac{N}{2}\left(\frac{1}{(N-1)t^2}-1\right)\lambda_0^2}\mathbb{1}\left(\lambda_0 \in \sqrt{\frac{N-1}{N}}G\right)\right].\end{aligned} \tag{10}$$

In this equation, λ_0 is the smallest eigenvalue of a random matrix from the GOE(N) ensemble.

We do not prove this lemma here, as it is proven as a particular case of Lemma 3.3. of [AAČ13], which is proven in the same paper. Applying this lemma to Eq. 9 with $t = \sqrt{\frac{p}{2(N-1)(p-1)}}$ and $G = (t\sqrt{N})B$ yields:

$$\mathbb{E}\,\mathrm{Crit}^0_{N,p}(B) = 2\sqrt{\frac{2}{p}}(p-1)^{\frac{N}{2}}\mathbb{E}_{\mathrm{GOE}(N)}\left[e^{-N\frac{p-2}{2p}\lambda_0^2}\mathbb{1}\left(\lambda_0 \in \sqrt{\frac{p}{2(p-1)}}B\right)\right] \tag{11}$$

2.4 The large N limit

We are interested in the annealed limit complexity $\lim_{N\to\infty}\frac{1}{N}\ln\mathbb{E}\,\mathrm{Crit}^0_{N,p}(B)$. It it thus clear from Eq. 11 that we need to study the large deviations for the smallest eigenvalue of a GOE(N) matrix. We state this result in an informal way, see Theorem A.1 of [AAČ13] for a more precise statement:

Lemma 2.3 (Informal) *Let λ_0 be the smallest eigenvalue of a GOE(N) matrix, and $I \subseteq \mathbb{R}$. Then*

$$\lim_{N\to\infty}\frac{1}{N}\ln\mathbb{P}\left[\lambda_0 \in I\right] = -\sup_{x\in I}F(x), \tag{12}$$

with F defined as:

$$F(x) \triangleq \begin{cases}\infty & \text{if } x \geq -\sqrt{2} \\ \frac{1}{\sqrt{2}}\int_{\sqrt{2}}^{-x}\mathrm{d}z\sqrt{\frac{z^2}{2}-1} & \text{otherwise}\end{cases}. \tag{13}$$

Using Lemma. 2.3 alongside Eq. 11 yields a final estimate for the annealed complexity:

$$\lim_{N\to\infty}\frac{1}{N}\ln\mathbb{E}\,\mathrm{Crit}^0_{N,p}(B) = \frac{1}{2}\ln(p-1) + \sup_{x\in\sqrt{\frac{p}{2(p-1)}}B}\left[-\frac{p-2}{2p}x^2 - F(x)\right]. \tag{14}$$

2.5 Description of the results

If one chooses in Eq. 14, $B = (-\infty, u)$ for $u \in \mathbb{R}$, one effectively counts the local minima of the p-spin of energy smaller than Nu. In this case, the supremum in Eq. 14 can be analytically performed, and one obtains an analytical form for $\lim_{N\to\infty}\frac{1}{N}\ln\mathbb{E}\,\mathrm{Crit}^0_{N,p}((-\infty,u))$. We can perform the same calculation we did for fixed points of any fixed index $k \in \mathbb{N}$, and we can define:

$$\Theta_k(u) \triangleq \lim_{N\to\infty}\frac{1}{N}\ln\mathbb{E}\,\mathrm{Crit}^k_{N,p}((-\infty,u)). \tag{15}$$

There are analytical expressions for all $\Theta_k(u)$ functions, see Eq.2.16 of [AAČ13]. Defining the *threshold* energy $E_\infty \equiv 2\sqrt{\frac{p-1}{p}}$, we can plot these functions, see Fig. 2.5.

Additional remarks

1. The minima always dominate for all energies below $-NE_\infty$, while for all energies above $-NE_\infty$, the complexity is dominated by critical points of diverging index.
2. One can perform similar calculations for critical points with indices that diverge with N, see for the rigorous derivation [AA+13].
3. This calculation can also be generalized to the mixed spherical p-spin case, see [AA+13], which exhaust all stationary isotropic Gaussian random fields on the sphere by Schoenberg's theorem.

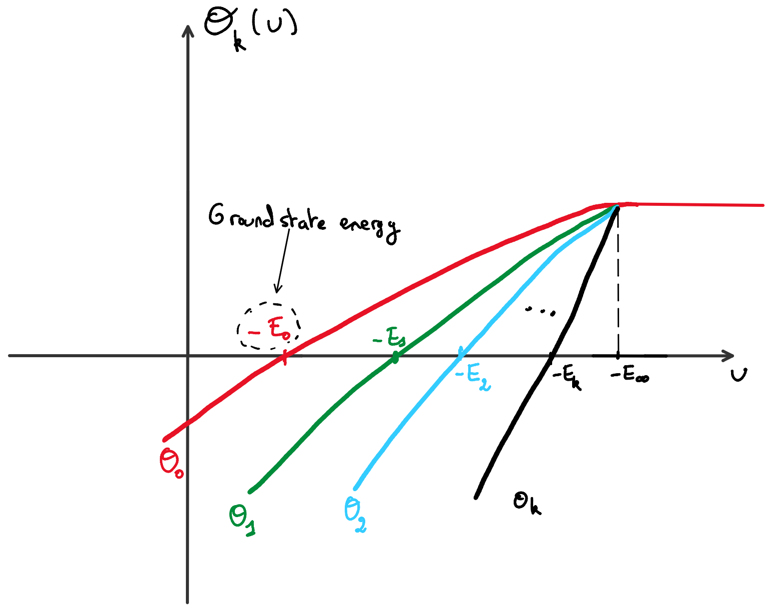

Figure 1: The functions Θ_k for the first indices

References

[AA+13] Antonio Auffinger, Gerard Ben Arous, et al. Complexity of random smooth functions on the high-dimensional sphere. *The Annals of Probability*, 41(6):4214–4247, 2013.

[AAČ13] Antonio Auffinger, Gérard Ben Arous, and Jiří Černỳ. Random matrices and complexity of spin glasses. *Communications on Pure and Applied Mathematics*, 66(2):165–201, 2013.

[AT09] Robert J Adler and Jonathan E Taylor. *Random fields and geometry*. Springer Science & Business Media, 2009.

[AW09] Jean-Marc Azaïs and Mario Wschebor. *Level sets and extrema of random processes and fields*. John Wiley & Sons, 2009.

[Fed59] Herbert Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–491, 1959.

[FW07] Yan V Fyodorov and Ian Williams. Replica symmetry breaking condition exposed by random matrix calculation of landscape complexity. *Journal of Statistical Physics*, 129(5-6):1081–1116, 2007.

[Fyo04] Yan V Fyodorov. Complexity of random energy landscapes, glass transition, and absolute value of the spectral determinant of random matrices. *Physical review letters*, 92(24):240601, 2004.

[MSWW63] John Willard Milnor, Michael Spivak, Robert Wells, and Robert Wells. *Morse theory*. Princeton university press, 1963.

[RABC19] Valentina Ros, Gerard Ben Arous, Giulio Biroli, and Chiara Cammarota. Complex energy landscapes in spiked-tensor and simple glassy models: Ruggedness, arrangements of local minima, and phase transitions. *Physical Review X*, 9(1):011003, 2019.

[S+17] Eliran Subag et al. The complexity of spherical p-spin models—a second moment approach. *The Annals of Probability*, 45(5):3385–3450, 2017.