Critical capacity of constrained perceptrons

1990 J. Phys. A: Math. Gen. 23 L847

(http://iopscience.iop.org/0305-4470/23/16/013)

J. Phys. A: Math. Gen. 23 (1990) L847-L851. Printed in the UK

LETTER TO THE EDITOR

Critical capacity of constrained perceptrons

Jorge Kurchan and Eytan Domany

Department of Electronics, Weizmann Institute of Science, Rehovot 76100, Israel

Received 25 May 1990

Abstract. We calculate the critical capacity of a single layer perceptron with interactions J_i constrained to take values on a general surface. The critical capacity is expressed in terms of a geometrical property of the constraining surface.

We present a calculation of the critical capacity of perceptrons with general constraints. Gardner [1] and Gardner and Derrida [2] calculated the critical capacity $\alpha_c(K)$ as a function of the embedding field strength K for the case when the vector of weights $\boldsymbol{J}$ can point in all directions; i.e. all normalized vectors that lie on a sphere centred at the origin are allowed. Subsequently, a number of calculations addressed the problem of capacity of perceptrons constrained in a variety of ways. For example, Amit *et al* [3] treated the problem of sign-constrained perceptrons, in which $J_i > 0$ for all i. Kanter and Eisenstein [4] generalized this to the case in which a fraction p of the weights is sign-constrained, whereas Kurchan and Domany [5] evaluated the capacity for $\boldsymbol{J}$ lying on a sphere whose centre is *shifted* from the origin. In this letter we show that all possible constraints can be treated in a unified manner, and the critical capacity can be cast in terms of an expression that reflects a geometrical property of the constraining surface or volume. Our result holds as long as capacity can be obtained from a replica-symmetric calculation.

We consider a perceptron whose vector of weights $\boldsymbol{J}$, of components J_j, is restricted to lie on a general surface:

$$|J|^2 = Nh^2(\boldsymbol{J}) \tag{1}$$

where h is a function of the direction $\boldsymbol{J}/|J|$, i.e.

$$h(\lambda\boldsymbol{J}) = h(\boldsymbol{J}) \qquad \sum_j \frac{\partial h}{\partial J_j} J_j = 0. \tag{2}$$

The unconstrained problem [1, 2] corresponds to $h = 1$. It is convenient to introduce another function of the direction, defined as $\beta(\boldsymbol{J}/|\boldsymbol{J}|) = 1$ if a line in the direction of $\boldsymbol{J}$ touches or pierces the surface, and $\beta(\boldsymbol{J}/|\boldsymbol{J}|) = 0$ otherwise.

For simplicity we consider here only such constraining surfaces that are met at most once by a ray emanating from the origin. Our treatment can be easily extended to more general situations by parametrizing each branch of the surface separately. We also assume $\boldsymbol{J}$ has been normalized so that $h = \mathrm{O}(1)$.

As in [1] we calculate the average of the logarithm of V, the J-space volume occupied by solutions that map αN independent random patterns $\boldsymbol{\xi}^\mu \equiv (\xi_1^\mu, \ldots, \xi_N^\mu)$ onto preassigned random $\xi^\mu = \pm 1$, with embedding strength K:

$$\overline{\ln V} = \sum_{\{\xi\}} \ln \int \Pi_\mu \Theta[(\boldsymbol{J} \cdot \boldsymbol{\xi}^\mu)\xi^\mu / N^{1/2} - K] \delta\left(\frac{|\boldsymbol{J}|^2}{h^2} - N\right) \beta(\boldsymbol{J}) \Omega(\boldsymbol{J}) \, \mathrm{d}^N J. \tag{3}$$

Here $\mathrm{d}^N J \equiv \Pi_j \, \mathrm{d}J_j$, and $\Omega(\boldsymbol{J})$ is a measure on the surface; it may contain, for example, the norm of the gradient of the argument of the delta function. We shall assume only that it is not exponential in N, i.e. that $(\ln \Omega)/N \to 0$ as $N \to \infty$.

We now change in (3) the integration variables from J_j to $\hat{J}_j = \boldsymbol{J}/h(\boldsymbol{J})$. The Jacobian of this transformation, easily obtained using (2), is given by $\mathrm{d}^N J = [h(\hat{J})]^N \, \mathrm{d}^N \hat{J}$ and the integral (3) becomes:

$$\overline{\ln V} = \sum_{\{\xi\}} \ln \int \Pi_\mu \Theta[(\hat{J} \cdot \boldsymbol{\xi}^\mu)\xi^\mu / N^{1/2} - K/h(\hat{J})] \delta(|\hat{J}|^2 - N) \beta(\hat{J}) \Omega(\hat{J}) h^N(\hat{J}) \, \mathrm{d}^N \hat{J}. \tag{4}$$

To calculate (4) we use the replica trick:

$$\overline{V^n} = \sum_{\{\xi\}} \int \Pi_{\mu,\alpha} \Theta[(\hat{J}_\alpha \cdot \boldsymbol{\xi}^\mu)\xi^\mu / N^{1/2} - K_\alpha] \delta(|\hat{J}_\alpha|^2 - N) \beta(\hat{J}_\alpha) \Omega(\hat{J}_\alpha) h^N(\hat{J}) \, \mathrm{d}^N \hat{J}_\alpha \tag{5}$$

with

$$K_\alpha = \frac{K}{h(\hat{J}_\alpha)}. \tag{6}$$

Following [1], we now express the Θ functions as integrals and sum their product over the ξ. This yields $\exp[\alpha N G_0]$, with

$$\exp[G_0(q_{\alpha\beta}, K_\alpha)] = \Pi_\alpha \int_{K_\alpha}^{\infty} \frac{\mathrm{d}\lambda_\alpha}{2\pi} \int_{-\infty}^{\infty} \Pi_\alpha \, \mathrm{d}x_\alpha \exp\left(i \sum_\alpha x_\alpha \lambda_\alpha - \frac{1}{2} \sum_{\alpha \neq \beta} q_{\alpha\beta} x_\alpha x_\beta - \frac{1}{2} \sum_\alpha x_\alpha^2 \right) \tag{7}$$

where

$$q_{\alpha\beta} = \frac{\hat{J}_\alpha \cdot \hat{J}_\beta}{N} \qquad \alpha \neq \beta. \tag{8}$$

We consider the case $K = K_\alpha = 0$ so that the function G_0 is the same as defined by Gardner in [1]. The integral (5) now reads:

$$\overline{V^n} = \int \Pi_\alpha [\delta(|\hat{J}_\alpha|^2 - N) \beta(\hat{J}_\alpha) \Omega(\hat{J}_\alpha) (h(\hat{J}_\alpha))^N \, \mathrm{d}^N \hat{J}_\alpha] \times \Pi_{\alpha\beta} \delta(1_{\alpha\beta} - \hat{J}_\alpha \cdot \hat{J}_\beta / N) \exp[\alpha N G_0(q_{\alpha\beta})] \Pi_{\alpha<\beta} \, \mathrm{d}q_{\alpha\beta}. \tag{9}$$

Exponentiating the second delta-function in the usual way [1], and assuming replica symmetry, this integral takes the form

$$\overline{V^n} = \exp\left[N\left(\alpha G_0(q) + \frac{n(1-n)}{2} F_q + G_1(F) \right) \right] \tag{10}$$

$$\mathrm{e}^{NG_1} = \int \Pi_\alpha [\delta(|\hat{J}_\alpha|^2 - N) \beta(\hat{J}_\alpha) \Omega(\hat{J}_\alpha) (h(\hat{J}_\alpha)^N \, \mathrm{d}^N \hat{J}_\alpha] \exp\left(\frac{F}{2} \sum_{\alpha \neq \beta} \hat{J}_\alpha \cdot \hat{J}_\beta \right). \tag{11}$$

In writing (10) we implicitly took the saddle point value of an integral over F and q.

We rewrite the last factor as:

$$\exp\left(\frac{F}{2}\sum_{\alpha\neq\beta}\hat{J}_\alpha\cdot\hat{J}_\beta\right)=\exp(-NnF/2)\int\frac{\mathrm{d}^N t}{(2\pi)^{N/2}}\exp\left(F^{1/2}\sum_\alpha \boldsymbol{t}\cdot\hat{J}_\alpha-\frac{|\boldsymbol{t}|^2}{2}\right) \tag{12}$$

where we have introduced an N-dimensional vector $\boldsymbol{t}$. Hence, replicas uncouple and we get for $n\to 0$

$$\frac{1}{nN}\ln\overline{V^n}=\operatorname{extr}\left(\alpha G_0(q)-\frac{F}{2}(1-q)+\int \Pi_j\,\mathrm{D}t_j\ln I\right) \tag{13}$$

with $\mathrm{D}t_j\equiv(2\pi)^{-1/2}\,\mathrm{e}^{-(t_j^2/2)}\,\mathrm{d}t_j$ and:

$$I^N(\boldsymbol{t})\equiv\int\exp[F^{1/2}(\boldsymbol{t}\cdot\hat{J})]\delta(|\hat{J}|^2-N)\beta(\hat{J})\Omega(\hat{J})(h(\hat{J}))^N\,\mathrm{d}^N\hat{J}. \tag{14}$$

The saddle point equation for F reads:

$$F^{1/2}(1-q)=\int\Pi_j\,\mathrm{D}t_j\frac{\langle\boldsymbol{t}\cdot\hat{J}\rangle}{N} \tag{15}$$

where we denote

$$\langle X\rangle=\frac{\int\mathrm{d}^N\hat{J}\,XW(\hat{J})}{\int\mathrm{d}^N\hat{J}\,W(\hat{J})} \tag{16}$$

with

$$W(\hat{J})=\exp[F^{1/2}(\boldsymbol{t}\cdot\hat{J})+N\ln h(\hat{J})]\delta(|\hat{J}|^2-N)\beta(\hat{J})\Omega(\hat{J}). \tag{17}$$

We now write $|\boldsymbol{t}|\equiv\sigma N^{1/2}$ and express the integral over $\boldsymbol{t}$ in spherical coordinates σ, ω_t (ω_t denotes all the angular variables associated with $\boldsymbol{t}$). Equation (15) becomes

$$F^{1/2}(1-q)=\left(\frac{N}{2\pi}\right)^{N/2}\int\mathrm{d}\omega_t\,\mathrm{d}\sigma\,\sigma^N\exp[-N\sigma^2/2]\langle\cos\gamma\rangle \tag{18}$$

where we introduced

$$\cos\gamma(\omega_t)\equiv\left[\frac{\boldsymbol{t}}{|\boldsymbol{t}|}\cdot\frac{\hat{J}}{N^{1/2}}\right]. \tag{19}$$

In order to calculate the critical capacity we let $q\to 1$. We first assume and later verify [1, 3] that in this limit $F\to\infty$. To calculate $\langle\cos\gamma\rangle$, note that it has in both numerator and denominator (see (15)-(17)) integrals that contain

$$\exp(N(F^{1/2}\sigma\cos\gamma+\ln h(\hat{J}))]. \tag{20}$$

when $F\to\infty$, the first term dominates the argument of the exponential, and the integrals in $\langle\cos\gamma\rangle$ are determined by a single value of $\hat{J}$, that which maximizes $\cos\gamma$ for fixed $\boldsymbol{t}$ over the allowed $\hat{J}$ values. Thus we get

$$F^{1/2}(1-q)=\left(\frac{N}{2\pi}\right)^{N/2}\int\mathrm{d}\sigma\,\sigma^N\exp(-N\sigma^2/2)\int\mathrm{d}\omega_t\,\max_{\hat{J}}(\cos\gamma). \tag{21}$$

As is evident already from this expression, the critical capacity depends only on the angular distribution of allowed $\boldsymbol{J}$ (and not on $|\boldsymbol{J}|$). As a trivial consequence of this, note that as long as the hypersurface of volume of allowed $\boldsymbol{J}$ encompasses the origin, one has $\alpha_c=2$ for $K=0$.

In the large-N limit the function $\sigma^N \exp(-N\sigma^2/2)$ is sharply peaked at $\sigma_0=1$. Evaluating (21) at the saddle point for σ, applying Stirling's formula and the expression $S_N=2\pi^{N/2}/\Gamma(N/2)$ for the surface of an N-dimensional unit sphere, we get for the limit $q\to 1$, $F\to\infty$

$$F^{1/2}(1-q)=\frac{1}{S_N}\int \cos\gamma_{\max}(\omega_t)\,\mathrm{d}\omega_t\equiv g \tag{22}$$

where

$$\cos\gamma_{\max}(\omega_t)\equiv \max_J\left[\frac{t}{t}\cdot\frac{\hat{J}}{N^{1/2}}\right] \qquad \text{for } \hat{J} \text{ such that } \beta(\hat{J})\neq 0 \tag{23}$$

Clearly, g is a purely geometric object that depends only on the angular distribution of the region of allowed $\hat{J}$. Since g is a positive number, we must have $F\to\infty$ when $q\to 1$, verifying our assumption.

The critical capacity α_c is obtained by substituting F from (22) into the saddle point equation for q:

$$\frac{\alpha_c}{g^2}\frac{\partial G_0}{\partial q}+\frac{1}{2(1-q^2)}=0. \tag{24}$$

This is the same saddle-point equation (for $q\to 1$) as that obtained by Gardner [1], but with α_c/g^2. Hence we get:

$$\alpha_c=2g^2. \tag{25}$$

In general this result holds only for $K=0$. However, if the constraint is such that $h(\boldsymbol{J})=\text{constant}$ (i.e. $\boldsymbol{J}$ is restricted to a subset of the surface of a sphere centred at the origin, as in [3] and [4]), the result

$$\alpha_c(K)=g^2\alpha_c^G(K) \tag{26}$$

(where $\alpha_c^G(K)$ is the capacity calculated in [1]), holds even for $K\neq 0$.

We demonstrate the usefulness of our general result (25) by calculating α_c for a few examples.

(a) In the trivial case of a surface that encompasses the origin, it is easy to see that $g=1$ and $\alpha_c=2$, as in [1].

(b) Consider the case in which the possible values of $\boldsymbol{J}$ lie on a sphere and are within a cone of angle ϕ (i.e. the maximum angle between two $\boldsymbol{J}$ is 2ϕ). The integral (22) defining g is divided in two parts: inside (a) and outside the cone. In the large-N limit this integral is dominated completely by the equator of the sphere that lies on the plane perpendicular to the cone's axis. This holds irrespective of the equator being inside or outside the cone. Therefore we get:

$$\begin{aligned} g&=1 \qquad \text{if } \phi>\pi/2\\ g&=\sin\phi \qquad \text{if } \phi<\pi/2. \end{aligned} \tag{27}$$

Since the condition for which our results hold for $K\neq 0$ is satisfied in this case, we have:

$$\begin{aligned} \alpha_c(K)&=\alpha_c^G(K) \qquad \text{if } \phi>\pi/2\\ \alpha_c(K)&=\alpha_c^G(K)\sin^2\phi \qquad \text{if } \phi<\pi/2. \end{aligned} \tag{28}$$

Since the fraction of the surface of the dome to the surface of the whole sphere is $\simeq\sin^N\phi$ we conclude that if this fraction is finite in the large-N limit, the capacity is α_c^G.

(*c*) Next we calculate the critical capacity for $\boldsymbol{J}$ constrained to a shifted sphere:

$$\sum_j (J_j - a)^2 = N. \tag{29}$$

For $a < 1$ clearly the origin is inside the allowed (shifted) sphere and, using (*a*) from above we have $g = 1$, $\alpha_c = 2$. For $a > 1$ denote by ϕ the angle between a line joining the origin with the centre of the sphere, and a line passing through the origin tangentially to the sphere. It is easy to see that $\sin\phi = 1/a$. Using this and (27) we get (valid for $K = 0$ only!) $g = 1/a$ and $\alpha_c = 2/a^2$ for $a > 1$. We have used (25) even though (for $a > 1$) the surface of the sphere is not of the type assumed in the derivation; nevertheless, the result is correct [5].

(*d*) We calculated expressions (22) and (25) for the case in which the J_j are constrained to a sphere and pN of them are constrained to have a definite sign [3, 4]. A rather lengthy calculation yields $g = (1 - p/2)^{1/2}$ and $\alpha_C(K) = (1 - p/2)\alpha_c^G(K)$.

(*e*) As a somewhat different application, it is easy to calculate (22) and (25) for the binary perceptron. In such a case one can show that one gets $g = (2/\pi)^{1/2}$ and $\alpha_c = 4/\pi$ which is the replica-symmetric result of [2] (in this case, however, replica symmetry is broken).

Finally let us remark that our derivation can be generalized in a number of ways. One straightforward generalization is the case in which the $\boldsymbol{J}$ are smoothly distributed on a *volume*, rather than a surface. In such a case the outer surfaces will dominate in the large-N limit. One such example is when the $\boldsymbol{J}$ take uniform values in a box that includes the origin (see, for example [6]). On one hand the outer surface (if it is reasonably smooth) dominates, but from the previous discussion we see that $\alpha_c = 2$ in all such cases.

Discussions with I Kanter are gratefully acknowledged. JK thanks the Campomar-Weizmann fund for a fellowship. ED thanks Minerva for partial support.

References

[1] Gardner E 1988 *J. Phys. A: Math. Gen.* **21** 257
[2] Gardner E and Derrida B 1988 *J. Phys. A: Math. Gen.* **21** 271
[3] Amit D, Wong K Y M and Campbell C 1989 *J. Phys. A: Math. Gen.* **22** 4687
[4] Kanter I and Eisenstein E 1990 *Preprint* Bar Ilan University
[5] Kurchan J and Domany E 1990 to be published
[6] Gutfreund H and Stein Y 1990 *J. Phys. A: Math. Gen.* **23** 2613