



Probability theory

Asymptotic variance and CLT for the number of zeros of Kostlan Shub Smale random polynomials



Variance asymptotique et TCL pour le nombre de zéros des polynômes aléatoires de Kostlan–Shub–Smale

Federico Dalmao ^{a,b}^a Departamento de Matemática y Estadística del Litoral, Universidad de la República, CP 50000, Salto, Uruguay^b Unité de Recherche en Mathématiques, Université du Luxembourg, campus Kirchberg, L-1359, Luxembourg

ARTICLE INFO

Article history:

Received 28 April 2015

Accepted after revision 17 September 2015

Available online 29 October 2015

Presented by the Editorial Board

ABSTRACT

In this note, we find the asymptotic main term of the variance of the number of roots of Kostlan–Shub–Smale random polynomials and prove a central limit theorem for this number of roots as the degree goes to infinity.

© 2015 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

R É S U M É

Dans cette note, nous calculons un équivalent de la variance asymptotique du nombre de racines réelles des polynômes aléatoires de Kostlan–Shub–Smale et nous démontrons un théorème de la limite centrale pour ce même nombre quand le degré tend vers l'infini.

© 2015 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

Consider the Kostlan–Shub–Smale (KSS for short) ensemble of random polynomials

$$X_d(x) := \sum_{n=0}^d a_n^{(d)} x^n; \quad x \in \mathbb{R},$$

where d is the degree of the polynomial and the coefficients $(a_n^{(d)})$ are independent centred Gaussian random variables such that $\text{var}(a_n^{(d)}) = \binom{d}{n}$.

Denote by N_d the number of real roots of X_d , that is

$$N_d := \#\{x \in \mathbb{R} : X_d(x) = 0\}.$$

It is well known that $\mathbb{E}(N_d) = \sqrt{d}$ [8,15]. The aim of this note is to prove the following result.

E-mail address: fdalmao@unorte.edu.uy.

Theorem 1.1. *The variance of the number of real roots N_d of KSS random polynomials verifies*

$$\lim_{d \rightarrow \infty} \frac{\text{var}(N_d)}{\sqrt{d}} = \sigma^2,$$

with $0 < \sigma^2 < \infty$ given in [Proposition 3.1](#) ($\sigma^2 \approx 0.57 \dots$). Furthermore,

$$\frac{N_d - \sqrt{d}}{d^{1/4}}$$

converges in distribution towards the centred normal distribution with variance σ^2 , $N(0, \sigma^2)$.

The number of roots of random polynomials has been under the attention of physicists and mathematicians for a long time. The first results for particular choices of the coefficients are due to Bloch and Pólya [5] in 1932. After many successive improvements and generalisations, in 1974 Maslova [12] stated the CLT for the number of zeros for polynomials with i.i.d. centred coefficients with finite variance. For related results, see the review by Bharucha–Reid and Sambandham [4] or the introduction in [3] and references therein.

The study of Kostlan–Shub–Smale ($m \times m$ systems) of polynomials started in the early nineties by Kostlan [8], Bogomolny, Bohigas and Lobœuf [6] and Shub and Smale [15]. The mean number of roots [8,15], some asymptotics as $m \rightarrow \infty$ for the variance [16] and for the probability of not having any zeros on given intervals (for $m = 1$) [14] are known. See also the review by Kostlan [9] and references therein.

We restrict our attention to the case $m = 1$. The mean number of real roots is \sqrt{d} [8,15]. This fact shows a remarkable difference with the polynomials with i.i.d. centred coefficients for which the asymptotic mean number of roots is $2 \log(d)/\pi$ [4].

Our tools are the Rice formulas for the (factorial) moments of the number of roots [2,7]; Kratz–León’s version of the chaotic expansion [10]; Kratz–León’s method for CLTs [11] and the Fourth Moment Theorem [13]. This method has been applied to the case of classical random trigonometric polynomials by Azaïs, Dalmao and León [3].

A key point in our analysis is that the covariance function of X_d , after a convenient rewriting, has a scaling limit. The limit covariance defines a centred stationary Gaussian process X on $[0, \infty)$. The asymptotic behaviour, as $d \rightarrow \infty$, of the number of real roots of X_d is intimately related to that of X on increasing intervals that eventually cover $[0, \infty)$. Similar situations occur in [1] and in [3]. The fact that in the present case the covariance of X and its spectral density are Gaussian is remarkable.

The note is organised as follows. Section 2 contains some preliminaries and sets the problem in a more convenient way. Section 3 presents the asymptotic behaviour of the variance of N_d . Section 4 contains the asymptotic normality of the standardised N_d and the proof of [Theorem 1.1](#).

2. Preliminaries

We start writing X_d in a more convenient way. Consider the polynomials

$$Y_d(t) := \sum_{n=0}^d a_n \cos^n(t) \sin^{d-n}(t), \quad t \in \mathbb{R}.$$

The polynomial Y_d is obtained from X_d after homogenising it as in [2]; restricting the domain to the unit circle S^1 and identifying a point in S^1 with a pair $(\sin(t), \cos(t))$. Note that x is a real root of X_d if and only if $t = \pm \arctan(x)$ are roots of Y_d . Hence, N_d coincides almost surely with the number of roots of Y_d on, say, $[0, \pi]$.

It is convenient to use the unit speed parameterisation. Let

$$Z_d(t) := Y_d\left(\frac{t}{\sqrt{d}}\right).$$

Denote by $N_{Z_d}([0, \sqrt{d}\pi])$ the number of roots of Z_d on $[0, \sqrt{d}\pi]$. Clearly,

$$N_d = N_{Z_d}\left([0, \sqrt{d}\pi]\right) \text{ almost surely.}$$

Direct computations show that Z_d is a centred stationary Gaussian process and that its covariance function r_d is given by

$$r_d(t) = \cos^d\left(\frac{t}{\sqrt{d}}\right). \tag{1}$$

Note that for $t \in [0, \sqrt{d}\pi/2]$ we have $r_d(\sqrt{d}\pi - t) = (-1)^d r_d(t)$. This fact implies that it suffices to deal with r_d restricted to $[0, \sqrt{d}\pi/2]$ as we will see in the sequel.

3. Asymptotic variance of the number of zeros

We need the following asymptotics and bounds.

Lemma 3.1. For each fixed $t \in \mathbb{R}$, we have

$$\cos^d \left(\frac{t}{\sqrt{d}} \right) \xrightarrow{d \rightarrow \infty} e^{-t^2/2}.$$

This convergence is uniform in t , for t being in a compact. Furthermore, for $0 < a < 1$ we have the following upper bounds

$$\cos^d \left(\frac{t}{\sqrt{d}} \right) \leq \begin{cases} e^{-\alpha t^2/2}; & \text{if } 0 \leq t < a\sqrt{d}, \\ \cos^d(a); & \text{if } a\sqrt{d} \leq t \leq \pi\sqrt{d}/2, \end{cases}$$

with $\alpha = 1 - a^2/3 \in (2/3, 1)$.

The proof of this lemma follows from Taylor–Lagrange expansion, to get the uniform convergence and the bounds it is necessary that cosines’ argument be less than 1.

Proposition 3.1 (Limit variance for N_d). As $d \rightarrow \infty$ we have

$$\frac{\text{var}(N_d)}{\sqrt{d}} \xrightarrow{d \rightarrow \infty} \sigma^2 := \frac{2}{\pi} \int_0^\infty \left(g(t) \left[\sqrt{1 - \rho^2(t)} + \rho(t) \arctan \left(\frac{\rho(t)}{\sqrt{1 - \rho^2(t)}} \right) \right] - 1 \right) dt + 1,$$

where g and ρ are given by

$$g(t) = \frac{1 - (1 + t^2) e^{-t^2}}{(1 - e^{-t^2})^{3/2}}, \quad \rho(t) = e^{-t^2/2} \frac{1 - t^2 - e^{-t^2}}{1 - e^{-t^2} - t^2 e^{-t^2}}.$$

Corollary 3.2. The asymptotic variance σ^2 is strictly positive.

Remark 1. The numerical approximation of this integral yields $\sigma^2 \approx 0.57 \dots$. Simulations are coherent with this value.

Sketch of the proof of Proposition 3.1. Recall that $\text{var}(N_d) = \mathbb{E}(N_d(N_d - 1)) - (\mathbb{E}(N_d))^2 + \mathbb{E}(N_d)$. From [7, Eq. 10.7.5], the second factorial moment of N_d verifies:

$$\frac{\mathbb{E}(N_d(N_d - 1))}{\sqrt{d}} = \frac{2}{\pi} \int_0^{\sqrt{d}\pi/2} g_d(t) \sqrt{1 - \rho_d^2(t)} dt + \frac{2}{\pi} \int_0^{\sqrt{d}\pi/2} g_d(t) \rho_d(t) \arctan \left(\frac{\rho_d(t)}{\sqrt{1 - \rho_d^2(t)}} \right) dt, \tag{2}$$

where

$$g_d(t) = \frac{1 - \cos^{2d} \left(\frac{t}{\sqrt{d}} \right) - d \cos^{2d-2} \left(\frac{t}{\sqrt{d}} \right) \sin^2 \left(\frac{t}{\sqrt{d}} \right)}{\left(1 - \cos^{2d} \left(\frac{t}{\sqrt{d}} \right) \right)^{3/2}},$$

$$\rho_d(t) = \cos^{d-2} \left(\frac{t}{\sqrt{d}} \right) \frac{1 - d \sin^2 \left(\frac{t}{\sqrt{d}} \right) - \cos^{2d} \left(\frac{t}{\sqrt{d}} \right)}{1 - \cos^{2d} \left(\frac{t}{\sqrt{d}} \right) - d \cos^{2d-2} \left(\frac{t}{\sqrt{d}} \right) \sin^2 \left(\frac{t}{\sqrt{d}} \right)}.$$

Now, from Lemma 3.1 is clear that, for fixed t we have $g_d(t) \rightarrow_d g(t)$ and $\rho_d(t) \rightarrow_d \rho(t)$.

In order to pass to the limit under the integral sign, we can obtain the domination by a careful analysis of Taylor–Lagrange expansions in a neighbourhood of 0. For the rest of the domain of integration, we use Lemma 3.1 to bound the integrand separately in the regions $0 \leq t < a\sqrt{d}$ and $a\sqrt{d} \leq t \leq \sqrt{d}\pi/2$. It follows that the contribution to the variance of the latter is negligible. Besides, note that the first integral in the r.h.s. of Equation (2) is only convergent after subtracting $(\mathbb{E}(N_d))^2/\sqrt{d} = \sqrt{d} = (2/\pi) \int_0^{\sqrt{d}\pi/2} dt$.

The result follows. \square

For the ease of presentation, the proof of Corollary 3.2 is postponed.

4. CLT for the number of roots

Proposition 4.1 (CLT). *As d tends to infinity,*

$$\frac{N_d - \mathbb{E}(N_d)}{d^{1/4}}$$

converges in distribution towards the centred normal distribution with variance σ^2 , $N(0, \sigma^2)$.

The sketch of the proof is as follows: we start with the chaotic expansion (3)–(4). Then, by the Kratz–León method [11], the finiteness of the variance of $N_d/d^{1/4}$ allows us to truncate the expansion and to derive its asymptotic normality from that of the sum of the first, say Q , terms. Finally, the Fourth Moment Theorem [13] gives a criterion to prove the asymptotic normality of the finite partial sums of the expansion.

Proof. We apply Kratz–León’s chaotic expansion [10] to the number of roots of Z_d on $[0, \sqrt{d}\pi]$. Hence,

$$\frac{N_d - \mathbb{E}(N_d)}{d^{1/4}} = \sum_{q=2}^{\infty} I_{q,d}^{Z_d} \tag{3}$$

where

$$I_{q,d}^{Z_d} = \frac{1}{d^{1/4}} \int_0^{\sqrt{d}\pi} f_q(Z_d(t), Z'_d(t)) dt, \quad f_q(x, y) = \sum_{\ell=0}^{\lfloor q/2 \rfloor} b_{q-2\ell} a_{2\ell} H_{q-2\ell}(x) H_{2\ell}(y), \tag{4}$$

with $a_{2\ell} = 2(-1)^{\ell+1} / (\sqrt{2\pi} 2^\ell \ell! (2\ell - 1))$, $b_k = \frac{1}{k!} \varphi(0) H_k(0)$ and H_k is the k -th Hermite polynomial. Note that we can delete the term corresponding to $q = 1$ since $H_1(0) = 0$.

We express $I_{q,d}^{Z_d}$ as multiple stochastic integrals w.r.t. a standard linear Brownian motion B . In the first place, we can write $Z_d(t) = \int_{\mathbb{R}} h_d(t, \lambda) dB(\lambda)$ with

$$h_d(t, \lambda) = \sum_{n=0}^d \sqrt{\binom{d}{n}} \cos^n\left(\frac{t}{\sqrt{d}}\right) \sin^{d-n}\left(\frac{t}{\sqrt{d}}\right) \mathbf{1}_{[n, n+1]}(\lambda). \tag{5}$$

Then, from Equation (5), using the properties of the chaos and the stochastic Fubini theorem, see [3, Remark 2], we have $I_{q,d}^{Z_d} = I_q^B(g_q(\lambda_q))$ with

$$g_q(\lambda_q) = \frac{1}{d^{1/4}} \int_0^{\sqrt{d}\pi} \sum_{j=0}^{\lfloor q/2 \rfloor} b_{q-2j} a_{2j} (h_d^{\otimes q-2j}(s, \lambda_{q-2j}) \otimes h_d'^{\otimes 2j}(s, \lambda_{2j})) ds;$$

where $\lambda_k \in \mathbb{R}^k$ and \otimes stands for the tensorial product.

By the properties of stochastic integrals, we have $I_q^B(g_q(\lambda_q)) = I_q^B(\tilde{g}_q(\lambda_q))$ being \tilde{g} the symmetrisation of g , that is, $\tilde{g}_q(\lambda_q) = \frac{1}{q!} \sum_{\sigma \in S_q} g_q(\lambda_\sigma)$, where S_q the group of all permutations of the set $\{1, \dots, q\}$ and $\lambda_\sigma = (\lambda_{\sigma(1)}, \dots, \lambda_{\sigma(q)})$.

Now, to get the asymptotic normality of the standardised zeros, by the Kratz–León method and the Fourth Moment Theorem, it suffices to prove that the contractions $\tilde{g}_q \otimes_k \tilde{g}_q(\lambda_{2q-2k})$ tend to 0 in L^2 as $d \rightarrow \infty$ for $q \geq 2$ and $k = 1, \dots, q - 1$ and $\lambda_{2q-2k} \in \mathbb{R}^{2q-2k}$. Let $\mathbf{z}_k = (z_1, \dots, z_k)$ and $\lambda_{2q-2k} = \lambda_{q-k} \otimes \lambda'_{q-k}$. The contractions are defined [13] as

$$\tilde{g}_q \otimes_k \tilde{g}_q(\lambda_{2q-2k}) = \int_{\mathbb{R}^k} \tilde{g}_q(\mathbf{z}_k, \lambda_{q-k}) \tilde{g}_q(\mathbf{z}_k, \lambda'_{q-k}) d\mathbf{z}_k.$$

The basic fact to compute the contractions is that the isometric property of stochastic integrals implies that $h_d^{\otimes p}(s) \otimes_k h_d^{\otimes p}(t) = r_d^k(t-s) h_d^{\otimes p-k}(s) \otimes h_d^{\otimes p-k}(t)$. Similarly, when the identified variable in the contraction involves the derivatives of h the result involves the derivatives of r_d . Taking this into account, it follows that

$$\|\tilde{g}_q \otimes_k \tilde{g}_q\|_2^2 = \frac{1}{d} \iiint \sum_{0 \leq j \leq \lfloor q/2 \rfloor} c_j \frac{1}{q!} \sum_{\sigma \in S_q} \prod_{i=0}^2 (r_d^{(i)}(t-s))^{\alpha_i} (r_d^{(i)}(t'-s'))^{\beta_i} (r_d^{(i)}(s-s'))^{\gamma_i} (r_d^{(i)}(t-t'))^{\delta_i} ds dt ds' dt';$$

where $\mathbf{j} = (j_1, j_2, j_3, j_4)$; vector inequalities are understood component-wise; $c_j = \prod_{i=1}^4 a_{2j_i} b_{q-2j_i}$; $\alpha_i = \alpha_i(\sigma, \mathbf{j})$, $\beta_i = \beta_i(\sigma, \mathbf{j})$, $\gamma_i = \gamma_i(\sigma, \mathbf{j})$ and $\delta_i = \delta_i(\sigma, \mathbf{j})$; $\sum_{i=1}^4 \alpha_i = \sum_{i=1}^4 \beta_i = k$ and $\sum_{i=1}^4 \gamma_i = \sum_{i=1}^4 \delta_i = q - k$. Actually, there are further

constrains for $\alpha, \beta, \gamma, \delta$ with respect to \mathbf{j} (namely $\alpha_1 \leq (q - 2j_1) \wedge (q - 2j_2)$, $\alpha_2 \leq (q - 2j_1) \wedge 2j_2 + (q - 2j_2) \wedge 2j_1$, etc.), but they are irrelevant for our purposes.

We bound the covariances by their absolute value. Since $\text{var}(Z_d(t)) = \text{var}(Z'_d(t)) = 1$, by the Cauchy–Schwarz inequality, each factor $|r_d^{(i)}(\cdot)| \leq 1$. Furthermore, since $k \geq 1$ and $q - k \geq 1$, we can bound from above the product of each group of factors (i.e.: with the same argument) by one of them. Hence, for some $i_1, i_2, i_3, i_4 \in \{0, 1, 2\}$, we have

$$\|\tilde{g}_q \otimes_k \tilde{g}_q\|_2^2 \leq \frac{C}{d} \iiint_{[0, \sqrt{d}\pi]^4} |r_d^{(i_1)}(t - s)r_d^{(i_2)}(t' - s')r_d^{(i_3)}(s - s')r_d^{(i_4)}(t - t')| \, ds \, dt \, ds' \, dt',$$

where C is a meaningless constant. Now, we make the change of variables: $(x, y, u, t') \mapsto (t - s, t' - s', s - s', t')$ and enlarge the domain of integration in order to have a rectangular one. Thus

$$\|\tilde{g}_q \otimes_k \tilde{g}_q\|_2^2 \leq \frac{C}{d} \int_0^{\sqrt{d}\pi} dt' \int_{-\sqrt{d}\pi}^{\sqrt{d}\pi} |r_d^{(i_1)}(x)| \, dx \int_{-\sqrt{d}\pi}^{\sqrt{d}\pi} |r_d^{(i_2)}(y)| \, dy \int_{-\sqrt{d}\pi}^{\sqrt{d}\pi} |r_d^{(i_3)}(u)| \, du.$$

Let us look at the three inner integrals. Note that since r_d is even, so is the absolute value of its derivatives, so it suffices to integrate on $[0, \sqrt{d}\pi]$. Besides, since for $t \in [0, \sqrt{d}\pi/2]$, we have $r_d(\sqrt{d}\pi - t) = (-1)^d r_d(t)$, it follows that $|r_d^{(i)}(\sqrt{d}\pi - t)| = |r_d^{(i)}(t)|$, $i \in \{0, 1, 2\}$. Then, we can further restrict the domain of integration to $[0, \sqrt{d}\pi/2]$. The finiteness of the integral then follows from Equation (1), by bounding the covariance by a polynomial (of degree at most 2) times $\cos^d(\cdot/\sqrt{d})$ and then using Lemma 3.1. Hence, the contractions tend to 0. The result follows. \square

Sketch of the proof of Corollary 3.2. We make use of the expansion of the number of zeros given by Equations (3)–(4). By the properties of the chaos, we have $\sigma^2 = \lim_d \sum_{q=2}^\infty \text{var}(I_{q,d}^{Z_d}) \geq \lim_d \text{var}(I_{2,d}^{Z_d})$. We can adapt easily Proposition 3.2 of [1] to obtain $\lim_d \text{var}(I_{2,d}^{Z_d}) = \lim_d \text{var}(I_{2,d}^X)$, where X is the centred stationary Gaussian process with covariance $r(t) = e^{-t^2/2}$ and $I_{2,d}^X$ is defined as in Equation (4) with Z_d replaced by X . In fact, Lemma 3.1 gives the limit covariance and Proposition 3.1 gives the domination. Actually, we need to restrict the domain of integration as above in order to use Lemma 3.1, we leave the details to the reader. Then, we can prove that $\lim_d \text{var}(I_{2,d}^X) > 0$ exactly as in [2, Eqs. 10.42–10.43]. The result follows. \square

Proof of Theorem 1.1. Put together Propositions 3.1 and 4.1 and Corollary 3.2. \square

References

[1] J.-M. Azaïs, J. León, CLT for crossings of random trigonometric polynomials, *Electron. J. Probab.* 18 (68) (2013) 68, 17 pp.
 [2] J.-M. Azaïs, M. Wschebor, *Level Sets and Extrema of Random Processes and Fields*, John Wiley & Sons Inc., Hoboken, NJ, ISBN 978-0-470-40933-6, 2009, xi+393 pp.
 [3] J.-M. Azaïs, F. Dalmao, J. León, CLT for the zeros of classical trigonometric polynomials, *Ann. Inst. Henri Poincaré* (2015), in press.
 [4] A. Bharucha-Reid, M. Sambandham, *Random Polynomials*. Probability and Mathematical Statistics, Academic Press, Inc., Orlando, FL, USA, ISBN 0-12-095710-8, 1986, xvi+206 pp.
 [5] A. Bloch, G. Pólya, On the roots of certain algebraic equations, *Proc. Lond. Math. Soc.* 33 (1) (1932) 102–114.
 [6] E. Bogomolny, O. Bohigas, P. Lebeuf, Distribution of roots of random polynomials, *Phys. Rev. Lett.* 68 (18) (1992) 2726–2729.
 [7] H. Cramér, M.R. Leadbetter, *Stationary and Related Stochastic Processes, Sample Function Properties and Their Applications*, reprint of the 1967 original, Dover Publications, Inc., Mineola, NY, USA, ISBN 0-486-43827-9, 2004, xiv+348 pp.
 [8] E. Kostlan, On the distribution of roots of random polynomials, in: *From Topology to Computation: Proceedings of the Smalefest*, Berkeley, CA, 1990, Springer, New York, 1993, pp. 419–431.
 [9] E. Kostlan, On the expected number of real roots of a system of random polynomial equations, in: *Foundations of Computational Mathematics*, Hong Kong, 2000, World Science Publishing, River Edge, NJ, USA, 2002, pp. 149–188.
 [10] M. Kratz, J. León, Hermite polynomial expansion for non-smooth functionals of stationary Gaussian processes: crossings and extremes, *Stoch. Process. Appl.* 66 (2) (1997) 237–252.
 [11] M. Kratz, J. León, Central limit theorems for level functionals of stationary Gaussian processes and fields, *J. Theor. Probab.* 14 (3) (2001) 639–672.
 [12] N. Maslova, The distribution of the number of real roots of random polynomials, *Teor. Veroâtn. Primen.* 19 (1974) 488–500.
 [13] G. Peccati, C. Tudor, Gaussian limits for vector-valued multiple stochastic integrals, in: *Séminaire de Probabilités XXXVIII*, in: *Lecture Notes in Math.*, vol. 1857, Springer, Berlin, 2005, pp. 247–262.
 [14] G. Schehr, S. Majumdar, Real roots of random polynomials and zero crossing properties of diffusion equation, *J. Stat. Phys.* 132 (2) (2008) 235–273.
 [15] M. Shub, S. Smale, Complexity of Bézout’s theorem. II. Volumes and probabilities, in: *Computational Algebraic Geometry*, Nice, 1992, in: *Progress in Mathematics*, vol. 109, Birkhäuser, Boston, MA, USA, 1993, pp. 267–285.
 [16] M. Wschebor, On the Kostlan–Shub–Smale model for random polynomial systems. Variance of the number of roots, *J. Complex.* 21 (6) (2005) 773–789.