

CUBIC FORMS IN GAUSSIAN VARIABLES

BY

F. ALBERTO GRÜNBAUM

Introduction

Let X_1, \dots, X_n be independent Gaussian variables with mean zero and variance one.

It is often of interest to determine the distribution of a polynomial in these variables, the simplest cases being the linear and the quadratic ones.

These two cases are fairly simple and can be best (for our purposes) summed up as follows. If $P(X_1, \dots, X_n)$ and $Q(X_1, \dots, X_n)$, in short $P(X)$ and $Q(X)$, have the same distribution, then there exists an orthogonal transformation O in R^n such that

$$P = Q \circ O. \quad (1)$$

In this paper we study the homogeneous cubic case. Our results are complete only in the case of a pair of Gaussian variables and thus we limit ourselves to this special case. The same methods, though, serve to uncover some results both for the cubic case in an arbitrary number of variables, as well as in the case of an arbitrary polynomial in two variables.

Preliminaries

We are interested in polynomials of the type

$$P(x, y) = ax^3 + bx^2y + cxy^2 + dy^3$$

and their distribution function when x and y are independent Gaussian variables with mean zero and variance one. This means that we are concerned with the function

$$\mu_P(\lambda) = \iint_{P_\lambda} e^{-(x^2+y^2)/2} dx dy \quad \text{where } P_\lambda = \{(x, y) \in R^2 \mid P(x, y) \leq \lambda\}.$$

The function $\mu_P(\lambda)$ is called the distribution of the polynomial P and can be defined in terms of the identity

$$\int_{R^2} e^{i\xi P(x,y)} e^{-(x^2+y^2)/2} dx dy = \int_{-\infty}^{\infty} e^{i\xi\lambda} d\mu_P(\lambda) \quad (2)$$

which holds for all real ξ .

For the cases when P is a linear or a quadratic polynomial (2) can be evaluated explicitly and one is led directly to conclude that there is a one-to-one relation

Received July 1, 1974.

between distribution functions and classes of polynomials equivalent under a rotation as in (1).

Beyond the quadratic case there is no hope of evaluating (2) explicitly and the analysis becomes more intricate.

The first thing to do is to parametrize classes of homogeneous cubic polynomials equivalent under a rotation in a convenient way. Write

$$P(x, y) = r^3[(\alpha e^{3i\theta} + \bar{\alpha}e^{-3i\theta}) + (\beta e^{i\theta} + \bar{\beta}e^{-i\theta})]$$

with α and β arbitrary complex numbers, and $x = r \cos \theta$, $y = r \sin \theta$.

It is convenient to set

$$A^2 = |\alpha|^2 + |\beta|^2$$

$$\alpha = A \cos R e^{i\theta_\alpha}, \quad \beta = A \sin R e^{i\theta_\beta}$$

$$T = \theta_\alpha - 3\theta_\beta.$$

Then, after an appropriate rotation in the plane $P(x, y)$ can be expressed as

$$P(x, y) = 2r^3 A(\cos R \cos(3\theta + T) + \sin R \cos \theta). \quad (3)$$

It is clear that if the set of all homogeneous cubic polynomials is partitioned into equivalence classes by the equivalence relation given in (1) then in each class we have one polynomial such as (3), and the parameters $A \geq 0$, $0 \leq R \leq \pi/2$, $0 \leq T \leq \pi$ serve to identify these classes in a one-to-one fashion.

We claim that the polynomials

$$P_1(x, y) = r^3 \cos \theta \quad \text{and} \quad P_3(x, y) = r^3 \cos 3\theta$$

have the same distribution function, i.e., $\mu_{P_1}(\lambda) = \mu_{P_3}(\lambda)$ while they are clearly not equivalent under a rotation. The contention above follows immediately from the identity

$$\int_0^{2\pi} e^{i\lambda \cos 3\theta} d\theta = \int_0^{2\pi} e^{i\lambda \cos \theta} d\theta$$

In terms of cartesian coordinates we have shown that the polynomials

$$P_1(x, y) = x^3 + xy^2 \quad \text{and} \quad P_3(x, y) = x^3 - 3xy^2$$

(plus those equivalent to these by means of rotations, and common scalar multiples of these) have the same distribution. The rest of the paper is devoted to showing that this is the only exceptional case. This is the content of the next theorem.

THEOREM. *Let P and Q be two homogeneous cubic polynomials. If their distribution functions coincide then either they are equivalent under an appropriate rotation $P = Q \circ O$, or else one is equivalent to $P_1 = Ar^3 \cos \theta$ and the other one to $P_3 = Ar^3 \cos 3\theta$.*

Proof of the theorem. Let $\mu_P(\lambda)$ be given. Denote by μ_n the “moments” of P , that is

$$\mu_n = \iint_{\mathbb{R}^2} P^n(x, y) e^{-(x^2+y^2)/2} dx dy = \int_{-\infty}^{\infty} \lambda^n d\mu_P(\lambda).$$

Notice that $\mu_{2n+1} = 0, n \geq 0$.

Recall now the parametrization of P given in (3). Notice that $\mu_2 = c_2 A^2$ with a universal constant c_2 , and thus A is known right away. Next we show that the knowledge of μ_4 and μ_6 suffices to decide the vanishing of $\sin R \cos R$. This product vanishes exactly if the polynomial in question is equivalent under a rotation to either $P_1 = Ar^3 \cos \theta$ or $P_3 = Ar^3 \cos 3\theta$. A straightforward computation gives

$$\mu_4 = c_4(3 + 6 \cos^2 R \sin^2 R + 4 \cos R \sin^3 R \cos T)$$

$$\mu_6 = c_6(1 + 6 \cos^2 R \sin^2 R + 3 \cos R \sin^3 R(\sin^2 R + 2 \cos^2 R) \cos T)$$

The constants involve only A but neither R nor T .

Considering μ_4 and μ_6 as functions of R and T , it is clear that they take the values $3c_4$ and c_6 , respectively, if $\cos R \sin R = 0$ or else if

$$\cos T = -\frac{3 \cos R}{2 \sin R} \quad (\text{from } \mu_4)$$

and

$$\cos T = -2 \frac{\cos R}{\sin R} \frac{1}{\sin^2 R + 2 \cos^2 R} \quad (\text{from } \mu_6).$$

These two equalities give

$$\frac{3 \cos R}{4 \sin R} = \frac{\cos R}{\sin R} \frac{1}{1 + \cos^2 R}$$

which has three possible solutions in $[0, \pi/2]$, namely $R = 0, \pi/2$ or $\cos^{-1} 1/\sqrt{3}$. But going back to the first equality with $R = \cos^{-1} 1/\sqrt{3}$ one gets $\cos T = -(3/2)(1/\sqrt{2})$ which rules out $\cos^{-1} 1/\sqrt{3}$.

Summing up, if we have $\mu_4 = 3c_4$ and $\mu_6 = c_6$ then $\sin R \cos R = 0$ and vice versa.

From now on we assume $\sin R \cos R \neq 0$.

Before proceeding it is convenient to observe that the knowledge of the distribution function of P is equivalent to that of the distribution function of P restricted to the unit circle. Indeed, observe that $P(r, \theta) = r^3 P(1, \theta)$ and thus

$$\int_0^\infty \int_0^{2\pi} P^n(r, \theta) e^{-r^2/2} dr d\theta = \left(\int_0^\infty e^{-r^2/2} r^{3n} dr \right) \int_0^{2\pi} P^n(1, \theta) d\theta$$

showing that the moments of $P(r, \theta)$ suffice to determine those of $P(1, \theta)$. Now since we are dealing with the finite interval $[0, 2\pi]$ these moments are enough to determine the distribution of $P(1, \theta)$.

Since A (see (3)) is already known, we consider for the rest of the paper that we know the distribution function of

$$P(\theta) = \cos R \cos (3\theta + T) + \sin R \cos \theta, \quad 0 \leq \theta \leq 2\pi \quad (4)$$

defined as

$$v(\lambda) = \text{Lebesgue measure of } \{\theta \mid P(\theta) \leq \lambda\}.$$

Determining R and T

The function $v(\lambda)$ introduced above grows steadily in the interval

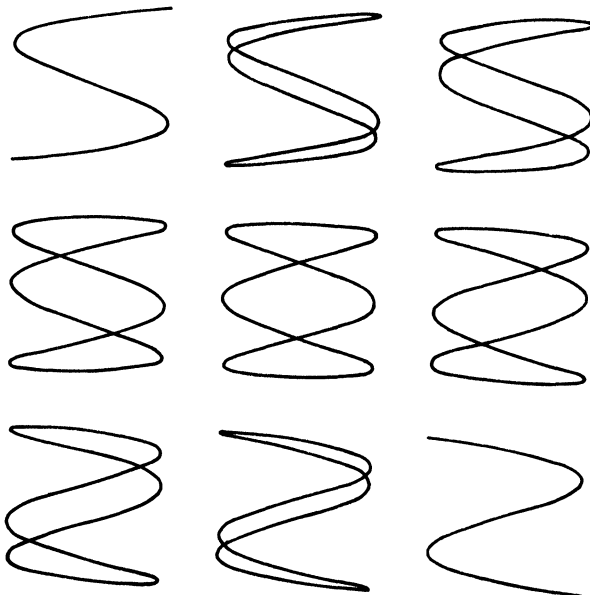
$$\min_{0 \leq \theta \leq 2\pi} P(\theta) \leq \lambda \leq \max_{0 \leq \theta \leq 2\pi} P(\theta).$$

In this section we aim at finding R and T , in (4), from the behavior of $v(\lambda)$ close to $\lambda = \max P(\theta)$. It is convenient to proceed by steps.

(I) An elementary analysis shows that the function $P(\theta)$ given in (4) takes its absolute maximum in $[0, 2\pi]$ only once, except in the case when $T = \pi$ and $9 \cos R > \sin R$.

Moreover one easily sees that, in the first case the maximum is taken at θ_0 with $(5/3)\pi \leq \theta_0 = \theta_0(R, T) \leq 2\pi$ while in the second case the maximum is taken again at $2\pi - \theta_0(R, \pi)$. For $T = \pi$ we have the symmetry $P(\pi + \theta) = P(\pi - \theta)$, $0 \leq \theta \leq \pi$.

These facts can be proved analytically quite simply. However a look at the accompanying Lissajous figures make them quite apparent. These figures can be used later on to suggest some further facts that will be needed. For $T = (k/8)\pi$, $k = 0, 1, \dots, 8$ we have plots of $x(\theta) = \cos (3\theta + T)$, $y(\theta) = \cos \theta$, $0 \leq \theta \leq 2\pi$.



For fixed values of R and T consider the family of straight lines

$$\cos Rx + \sin Ry = \text{constant.} \tag{5}$$

The maximum value of $P(\theta)$ corresponds to that straight line with largest possible value of (5), which intersects the Lissajous figure in question. A look at the figures shows that this gives a single value of θ_0 if $T \neq \pi$. For $T = \pi$ this is the case exactly if $-(\cos R)/(\sin R) \geq -(1/9)$.

(II) Denote with θ_0 the (unknown) value of θ where the absolute maximum of $P(\theta)$ is taken. In the case of $T = \pi$, take $(5/3)\pi \leq \theta_0 \leq 2\pi$. In the neighborhood of θ_0 , P has a power series expansion

$$P(\theta) = \alpha + \sum_{n \geq 2} \alpha_n (\theta - \theta_0)^n, \quad \alpha_2 < 0, \quad \alpha = \max_{0 \leq \theta \leq 2\pi} P(\theta).$$

Locally we get two power series in $\xi = P - \alpha$ by inverting (6), namely

$$\theta - \theta_0 = \sum_{n \geq 1} C_{2n} \xi^{2n/2} \pm \sum_{n \geq 1} C_{2n-1} \xi^{(2n-1)/2} \tag{7}$$

The coefficients C_n are determined in a straightforward fashion, the first few equations are

$$\begin{aligned} \alpha_2 C_1^2 &= -1 \\ 2\alpha_2 C_1 C_2 + \alpha_3 C_1^3 &= 0 \\ \alpha_2 (C_2^2 + 2C_1 C_3) + 3\alpha_3 C_1^2 C_2 + \alpha_4 C_1^4 &= 0. \end{aligned} \tag{8}$$

It follows easily that the C_{2n} are unambiguously determined while the C_{2n+1} are determined up to a common change in sign.

The distribution function of $P(\theta)$, $\nu(\lambda)$, gives us some information about the coefficients introduced above. To the left of $\lambda = \alpha$ the expansion

$$2\pi - \nu(\lambda) = \sum_{n \geq 1} b_{2n-1} (\alpha - \lambda)^{(2n-1)/2}, \quad b_1 > 0 \tag{9}$$

holds for λ close enough to α .

From the observations made earlier, it is now clear that if the absolute maximum of $P(\theta)$ is taken only once in $0 \leq \theta \leq 2\pi$, $b_{2n-1} = 2C_{2n-1}$ while if this absolute maximum is taken twice $b_{2n-1} = 4C_{2n-1}$. We postpone an argument showing how to decide, from $\nu(\lambda)$ alone which of the two cases hold. Taking this point for granted we can now compute α and the coefficients C_{2n-1} from $\nu(\lambda)$. From (8) we can find α_2 , but not α_3 .

Now notice that

$$\begin{aligned} \alpha &= \cos R \cos (3\theta_0 + T) + \sin R \cos \theta_0 \\ 2\alpha_2 &= -9 \cos R \cos (3\theta_0 + T) - \sin R \cos \theta_0 \end{aligned} \tag{10}$$

and we can solve for $\cos R \cos (3\theta_0 + T)$ and $\sin R \cos \theta_0$. Then all even-order derivatives of $P(\theta)$ at θ_0 are now known, i.e., we can determine all the

coefficients α_{2n} . From (8) we get (note that $C_1 \neq 0$), $C_2 = -(\alpha_3 C_1^2)/(2\alpha_2)$ and substituting in the third equation in (8) we get

$$\frac{7}{4} \frac{C_1^4}{\alpha_2} \alpha_3^2 + 2C_1 C_3 \alpha_2 + \alpha_4 C_1^4 = 0$$

and thus while α_3 is not known directly we have α_3^2 . But this does the job, since the first and third derivatives of $P(\theta)$ at θ_0 are

$$3 \cos R \sin (3\theta_0 + T) + \sin R \sin \theta_0 = 0$$

$$27 \cos R \sin (3\theta_0 + T) + \sin R \sin \theta_0 = 6\alpha_3$$

which give $\alpha_3 = -(4/3) \sin R \sin \theta_0$. But $(5/3)\pi \leq \theta_0 \leq 2\pi$, and α_3 is thus nonnegative. Summing up, we have determined the four quantities

$$\sin R \cos \theta_0, \quad \sin R \sin \theta_0, \quad \cos R \cos (3\theta_0 + T), \quad \cos R \sin (3\theta_0 + T).$$

The ratio of the first two of them (recall that $R \neq 0, \pi/2$), and the condition $(5/3)\pi \leq \theta_0 \leq 2\pi$ gives θ_0 , and a fortiori R . It is now trivial to get T from $\cos (3\theta_0 + T)$ and $\sin (3\theta_0 + T)$.

(III) In this last step we show how to distinguish from $v(\lambda)$ alone if the absolute maximum of $P(\theta)$ is taken once or twice. At this point the Lissajous figures become quite suggestive.

Start with the observation that $v(\lambda)$, which satisfies $v(\lambda) + v(-\lambda) \equiv 2\pi$, can be

- (a) differentiable for $-\alpha < \lambda < \alpha$;
- (b) not differentiable for a single value $0 \leq \lambda_1 < \alpha$;
- (c) not differentiable for more than two values of $-\alpha < \lambda < \alpha$.

Case (c) calls for $T \neq \pi$ and thus $P(\theta)$ takes its maximum only once. The other two cases will be analyzed separately. If (a) holds then we have $T = \pi$, $9 \cos R > \sin R$ only if $\cos R = \sin R = 1/\sqrt{2}$, which leads to

$$\alpha = \frac{4}{3} \sqrt{\frac{2}{3}} \quad \text{and} \quad \alpha_2 = -\frac{8}{\sqrt{6}}.$$

Now from $v(\lambda)$ one reads off α and b_1 , see (9). $T = \pi$, $9 \cos R > \sin R$ implies $\alpha = (4/3)\sqrt{(2/3)}$ and $b_1^2 = 2\sqrt{6}$. If these are the values obtained from $v(\lambda)$ and the maximum was taken only once we would get $\alpha_2 = -(2/\sqrt{6})$. But the pair

$$(\alpha, \alpha_2) = \left(\frac{4}{3} \sqrt{\frac{2}{3}}, -\frac{2}{\sqrt{6}} \right)$$

leads to a contradiction since (10) gives

$$\sin R \cos \theta_0 = \frac{9\alpha}{8} + \frac{\alpha_2}{4}$$

and we would have $9\alpha/8 + \alpha_2/4 = \sqrt{(50/48)} > 1$. In brief if (a) holds the case $T = \pi$, $9 \cos R > \sin R$ is exactly equivalent to

$$\alpha = \frac{4}{3} \sqrt{\frac{2}{3}} \quad \text{and} \quad b_1^2 = 2\sqrt{6}.$$

If (b) holds and $T \neq \pi$ the left derivative of $v(\lambda)$ at λ_1 is larger than the right derivative at this point. If $T = \pi$, and $-1 < -(\cos R/\sin R) < -(1/9)$ the right derivative at λ_1 exceeds the left one, while if $-(\cos R/\sin R) < -1$ then the left derivative at λ_1 exceeds the right one. On account of this, if (b) holds it is enough that we learn how to separate the cases $T \neq \pi$ and $T = \pi$,

$$-\frac{\cos R}{\sin R} < -1.$$

In the first case consider α and λ , as functions of $0 \leq T \leq \pi$. $\alpha(T)$ is strictly decreasing while $\lambda_1(T)$ increases strictly as T goes from 0 to π , and we have

$$\lim_{T \rightarrow \pi} \alpha(T) = \lim_{T \rightarrow \pi} \lambda_1(T) = \sqrt{\frac{32}{27}}.$$

In the second case, $T = \pi$, consider α and λ , as functions of $-\infty < -(\cos R/\sin R) < -1$. We get

$$\alpha = \frac{2}{3} \frac{(3 \cos R + \sin R)^{3/2}}{(12 \cos R)^{1/2}}, \quad \lambda_1 = \cos R - \sin R.$$

α is first increasing and then it decreases to the value $\sqrt{(32/27)}$ for $-(\cos R/\sin R) = -1$. λ_1 steadily decreases from 1 to 0.

The information given above can be used in an elementary analysis to prove that if (b) holds an inspection of α and λ_1 allows to decide between $T \neq \pi$ and $T = \pi$, $-(\cos R/\sin R) < -(1/9)$. The proof of the theorem is now complete.

UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA