HAL
open science

Higher Order variance and Gauss Jacobi Quadrature : II

René Blacher

▶ To cite this version:

René Blacher. Higher Order variance and Gauss Jacobi Quadrature : II. [Research Report] LJK. 2012. hal-00736766

HAL Id: hal-00736766
https://hal.science/hal-00736766v1

Submitted on 29 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Higher Order variance and Gauss Jacobi Quadrature : II

René BLACHER

Laboratory LJK
Université Joseph Fourier
Grenoble
France

Summary : In this report, we continue the study of higher order variances and quadrature Gauss Jacobi. Recall that the variance of order j measures the concentration of a probability close to j points $x_{j,s}$ with weight $\lambda_{j,s}$ which are determined by the parameters of the quadrature Gauss Jacobi. We shall study the normalized variances. We shall specify how to detect the parameters of Gaussian mixtures

Summary : Dans ce rapport, nous poursuivons l'étude des variances d'ordre supérieur et la quadrature de Gauss Jacobi. On rappelle que la variance d'ordre j mesure la concentration d'une probabilité autour de j points $x_{j,s}$ avec des poids $\lambda_{j,s}$ qui sont déterminés par les paramêtres de la quadrature de Gauss Jacobi. On étudiera le comportement des variances normalisées et on précisera comment détecter les paramêtres des mélanges gaussiens.

Key Words : Higher order variance, Gauss Jacobi quadrature, Central limit theorem, Higher order regression. Gaussian mixtures.

Contents

1 Higher Order Variances **3**
1.1 Introduction . . . 3
1.2 Some examples . . . 4
1.2.1 Bienayme-Tschbichev Inequality . . . 4
1.2.2 Examples . . . 5
1.2.3 Theoretical examples . . . 6
1.3 Some properties . . . 7

2 Estimation **13**

3 Detection of points of concentration **19**
3.1 Introduction . . . 19
3.2 Example 1 . . . 19
3.3 Example 2 . . . 22

4 Application : mixtures **23**
4.1 Some properties . . . 23
4.2 First application to mixtures . . . 24
4.2.1 Method . . . 24
4.3 Second application to mixtures . . . 26
4.3.1 Presentation . . . 26
4.3.2 Example . . . 26
4.3.3 Calculation of the first standard deviation . . . 26
4.3.4 Deletion of the first Gaussian component . . . 28
4.3.5 Estimation of the Other Gaussian components . . . 29
4.3.6 Conclusion . . . 29

A Standardized Variances **30**
A.1 Centered Gaussian distributionss . . . 30
A.2 Gaussian mixtures . . . 30
A.3 Truncated gaussian mixtures . . . 37
A.3.1 First type of graphs . . . 38
A.3.2 Second type of graphs . . . 40
A.4 Comparison of the various results . . . 45
A.5 Conclusion . . . 47

B Example of the decomposition of a Gaussian mixture **48**
B.1 Study of the case $s = 0.4472$ . . . 48
B.2 Study of the case $s = 0.42$ . . . 49
B.3 Study of the case $s = 0.465$ . . . 51
B.4 Conclusion . . . 52

Chapter 1

Higher Order Variances

1.1 Introduction

Orthogonal polynomials have many interesting applications in Probability and Statistics. So they have introduced higher order correlation coefficients and higher order variances (cf [3], [2], [5], [6], [8], [7], [4]). They also have introduced new hypotheses for the central limit theorem (cf [4]). They allow also to obtain the distributions of quadratic forms, Gaussian or not Gaussian, and simple methods of calculation of these laws (cf [10]).

Higher order variances have been introduced in [7] and [8]. They generalize the classical variance. Thus, variance of order 1 measures the concentration of a probability close to a point: the expectation. Variance of order j measures the concentration close to j points which are the zeros of the j-th orthogonal polynomial.

Notations 1.1.1 *Let X be a random variable defined on (Ω, A, P). Let m be the distribution of X. Let $\tilde{P}_j$ be the j-th orthogonal polynomial associated to X such that $\tilde{P}_j(x) = \sum_{t=0}^{j-1} a_{j,t}x^t + x^j$.*

We set $n_0^m = dim\big\{L^2(\mathbb{R}, m)\big\}$. Let $\Theta \subset \mathbb{N}$ such that $\tilde{P}_j(x)$ exists if $j \in \Theta$. We denote by P_j the j-th orthonormal polynomial associated to X if there exists.

Remark that if m is concentrated in n_0^m points when $n_0^m < \infty$, then $\Theta = \{0, 1, ..., n_0^m\}$. If not, $\Theta = \mathbb{N}$ if all moments exists, and $\Theta = \{0, 1, ..., d\}$ if $\int |x|^{2d-1}.m(dx) < \infty$ and $\int |x|^{2d+1}.m(dx) = \infty$. In this case, P_j exist if $\int |x|^{2d}.m(dx) < \infty$.

For example, $\tilde{P}_0 \equiv 1$, $\tilde{P}_1(x) = x - \mathbb{E}\{X\}$, $\tilde{P}_2(x) = x^2 - \frac{M_3 - M_1M_2}{M_2 - M_1^2}(x - M_1) - M_2$, where $M_s = \mathbb{E}\{X^s\}$ and where $\mathbb{E}\{.\}$ is the expectation.

Now we know that the zeros of $\tilde{P}_j$ are real (cf th 5-2 page 27 [13])

Proposition 1.1.1 *Let $j \in \Theta$. Then, the zeros of $\tilde{P}_j$ are distincts and real. We denote them by $x_{j,s}$, s=1,2,....,j.*

For example, $x_{1,1} = \mathbb{E}\{X\}$, $x_{2,s} = \frac{M_3 - M_1M_2}{2(M_2 - M_1^2)} \pm \frac{1}{2}\sqrt{\left(\frac{M_3 - M_1M_2}{M_2 - M_1^2}\right)^2 - 4M_2}$.

Now the zeros of orthogonal polynomials have stronger properties : in particular the Gauss-Jacobi Quadrature.

Theorem 1 *Let $j \in \Theta$. There exists a single probability m^j concentrated over j distincts points such that $\int x^q.m(dx) = \int x^q.m^j(dx)$ for q=0,1,...,2j-1.*

Moreover, the j points of concentration of m^j are the j zeros of $\tilde{P}_j$: $x_{j,s}$, s=1,...,j, and the probabilities $\lambda_{j,s} = m^j\big(\{x_{j,s}\}\big)$ are the Christoffel numbers which check $\lambda_{j,s} = \int \ell_s^j(x).m(dx)$, where $\ell_s^j(x) = \frac{\tilde{P}_j(x)}{(x - x_{j,s})\tilde{P}_j'(x_{j,s})}$ when $\tilde{P}_j'$ is the derivative of $\tilde{P}_j$.

The proof is in [13] if $j \neq n_0^m$. If $j = n_0^m$, it is in [8] (cf also [9]).

For example, if j=2, $\ell_1^2(x) = \frac{x-x_{2,2}}{x_{2,1}-x_{2,2}}$ and $\ell_2^2(x) = \frac{x-x_{2,1}}{x_{2,2}-x_{2,1}}$. Therefore, $\lambda_{2,1} = \frac{M_1-x_{2,2}}{x_{2,1}-x_{2,2}}$ and $\lambda_{2,2} = \frac{M_1-x_{2,1}}{x_{2,2}-x_{2,1}}$.

Now, we complete the definition of Gauss Jacobi quadrature by defining higher order variances.

Definition 1.1.2 *Let $j \in \Theta$. We call variance of order j, and we note it by σ_j^2 or $\sigma_j(X)^2$ or $\sigma_j(m)^2$ the real $\sigma_j^2 = \int |\tilde{P}_j|^2.dm$.*

Remark that $\tilde{P}_j = \sigma_j P_j$. Moreover, $\sigma_1(X)^2 = M_2 - M_1^2$ is the classical variance. Now, if j=2, $\sigma_2^2 = M_4 - \frac{(M_3-M_1M_2)^2}{M_2-M_1^2} - M_2^2$.

Then, the variance of order j measure the concentration of X close to j distinct points.

Theorem 2 *Let $j \in \Theta$. Then, $\sigma_j = 0$ if and only if m is concentrated in j distincts points which are the zeros of $\tilde{P}_j$: the $x_{j,t}$'s. Moreover the probability associated at each $x_{j,t}$ is equal at $\lambda_{j,t}$. In this case, $j = n_0^m < \infty$ and $\tilde{P}_j = 0$ in $L^2(\mathbb{R}, m)$.*

Then, variances of higher order generalize the classical variance which one can call variance of order 1. Indeed, classical variance measure the concentration close to expectation. For the variance of order j, the zeros of $\tilde{P}_j$ plays this role. Moreover we know the weight associated : the $\lambda_{j,t}$'s. All these properties justify well the name of variances of higher order.

We wonder then which applications are possible. We think immediately to Gaussians mixtures. We are going to recall in this report that we can indeed apply these variances to solve this problem.

1.2 Some examples

In this section, we shall study some examples which allow to well understand the role which can play higher order variances and Gauss Jacobi quadrature in the distribution of probability.

1.2.1 Bienayme-Tschbichev Inequality

At first, recall that the Bienayme-Tschbichev Inequality allows to specify more this concentration autour de j points distincts.

Proposition 1.2.1 *Let $\epsilon > 0$. Then, $P\big(|\tilde{P}_j| > \epsilon\big) \leq \frac{\sigma_j^2}{\epsilon^2}$.*

In particular assume that σ_j^2 is small enough. Let ω such that $|\tilde{P}_j(X(\omega))| \leq \epsilon$. Then, with a strong probability, there exists s such that $|X(\omega) - x_{j,s}| \leq \epsilon$. Then, the variance of order j measures the concentration of a probability close to j distincts points.

For example, let us study $\tilde{P}_3(x) = (x-0.95)(x-0.5)(x-0.1)$: cf figure 1.1. In figure 1.1, we see the points x such that $|\tilde{P}_3(x)| = |(x-0.95)(x-0.5)(x-1)| \leq 0.01$: they are the points of real axis which belong to intervals I_1, I_2, I_3 intersection of the 3 rectangles of the figure and of the real axis. Then, $P\{X \notin I_1 \cup I_2 \cup I_3\} = P\big\{|\tilde{P}_j(x)| > 0.01\big\} \leq 10^4\sigma_j^2 = 0.1$ if $\sigma_j^2 = 10^{-5}$.

Then, if σ_3^2 is small (here 10^{-5}) $P\{X \notin I_1 \cup I_2 \cup I_3\}$ is small also, i.e. X is concentrated close to points $x_{3,1} = 0.1$, $x_{3,2} = 0.5$, $x_{3,3} = 0.95$.

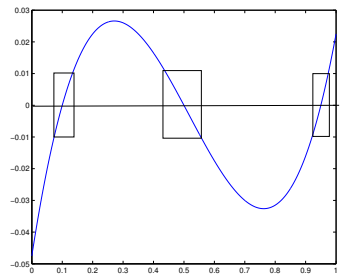

Figure 1.1: $\tilde{P}_3(x) = (x - 0.95)(x - 0.5)(x - 0.1)$

We also notice that what teaches us the Bienayme Tchebicheff inequality is generally enough limited. Indeed, it is necessary that σ_3^2 is very small in order to be sure that X is concentrated close to $x_{3,1}$, $x_{3,2}$, $x_{3,3}$: It is not surprising because this inequality is rather unrefined.

1.2.2 Examples

Indeed, when we study examples, we understand that higher order variances give better indications than it. In a enough large number of case where the probability are rather concentrated around certain points, we obtain results corresponding better to what we can hope.

Indeed, the following figures are clear enough to get an idea of density and weight $\lambda_{j,t}$'s of various probabilities.

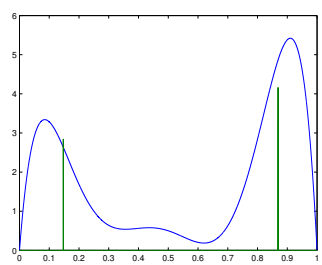

Figure 1.2: $x_{2,t}$ =0.8691, 0.1473, $\lambda_{2,t}$= 0.5944, 0.4056, σ_2^2=0.0037

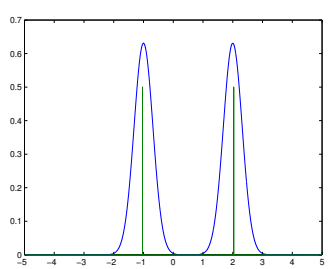

Figure 1.3: Gaussian mixtures : $x_{2,t}$=2.0330, -1.0330, $\lambda_{2,t}$=0.5000, 0.5000, σ_2^2=0.9200

Many other examples confirm these results : cf [9] section1.1.1 (cf also section A).

Now, we see that concentrations of figures 1.2.2 and 1.2.2 are not so different. Nevertheless the variances are completement different. It is due to the fact that figure 1.2.2 is concentrated in

[0,1] and figure 1.2.2 is concentrated in [-5,5] (cf appendix A).

Then, one asks oneself if there exists a way for normalizing variances. Then, a normalization can may be done by considering the number $\hat{\sigma}_j = \frac{\sigma_j}{||x^j||}$ which represents the sinus of the angle between the polynomial x^j and the subspace spanned by polynomials of degree strictly less than j.

In appendix A, we understand that it is a good standardization. In particular, if we apply an homothety, normalized variances do not change (cf proposition 1.3.8).

Unfortunately, it is insufficient because if a translation is done, variances do not change (cf proposition 1.3.3), but standardized variances change. Finally if we want to standardize well the variances, it seems that it is necessary to use at first the standardized variances $\hat{\sigma}_j^2$, and then, or to center the random variable X, or to impose by translation that it is positive with a probability very strong.

It seems that it is the good method. In that case, only in the value of the standardized variance of order j, we shall know generally if there is concentration or not near j distincts points.

1.2.3 Theoretical examples

Now, we can also calculate the variances of the classic distributions. Here we give this result for the normal laws. But we can find the variances of the other classic laws page 21 of [9].

Proposition 1.2.2 *Let $\hat{H}_j(x) = e^{x^2} \frac{d^j(e^{-x^2})}{dx}$ be a Hermite orthogonal polynomial. We suppose that X has the $N(m, \sigma^2)$ distribution. We denote by $H_j^{m\sigma^2}$ and $\left(\sigma_j^{m\sigma^2}\right)^2$ the associated Hermite orthonormal polynomials of order j and variances of order j. Then,*

$$H_j^{m\sigma^2}(x) = \frac{(-1)^j \sigma^j}{2^{j/2}} \hat{H}_j\Big(\frac{x-m}{\sigma\sqrt{2}}\Big) \ ,$$

$$\left(\sigma_j^{m\sigma^2}\right)^2 = j!\sigma^{2j} \ .$$

For example for the distribution N(0,0.1), one can look at figure 1.2.3

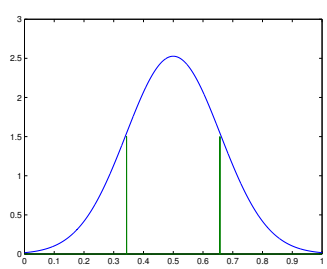

Figure 1.4: Distribution N(0,0.1), $x_{2,t}$=0.6568, 0.3433, $\lambda_{2,t}$=0.5000, 0.5000, σ_2^2=0.0011

Let us remark that

$$\left(\hat{\sigma}_j^{0\sigma^2}\right)^2 = \frac{j!\sigma^{2j}}{\frac{2j!\sigma^{2j}}{2^j j!}} = \frac{2^j (j!)^2}{2j!} \ .$$

Then, for j=1,2,3,4, $\left(\hat{\sigma}_j^{0\sigma^2}\right)^2 \approx$ 1, 0.666, 0.4, 0.228.

Then, even if σ_j^2 is small, σ_j^2 can mean not that there is a concentration close to j distinct points. It is enough that the classical variance of a Gaussian distribution is small.

Remark that, generally $\frac{2^j (j!)^2}{2j!}$ is bigger than the $\hat{\sigma}_i^2$'s of the appendix A. Then, the $\hat{\sigma}_j^2$'s seems well to be good indicators of the concentration of a probability near j distinct points. Let us notice that, in that case, we consider gaussian centered variables, what is one of means of standardization of the variances which we envisage.

1.3 Some properties

Points of concentration of a probability can be detected using various properties of the Gauss Jacobi Quadrature. The most important of these properties is the Stieltjes-Markov Inequality.

Proposition 1.3.1 *Let F_X be the distribution function of X. Then, for all k ∈1,2,..,j,*

$$\sum_{x_{j,s}<x_{j,k}} \lambda_{j,s} \leq F_X(x_{j,k}-0) \quad \textit{and} \quad \sum_{x_{j,s}\leq x_{j,k}} \lambda_{j,s} \geq F_X(x_{j,k}+0) \ .$$

These results are proved pages 26-29 of [14] : equation 5.4. For example, in figure 1.5, we have the distribution function of m and m^j.

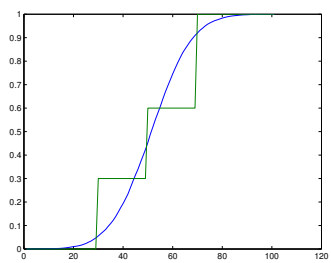

Figure 1.5: Stieljes-Markov Inequality

In particular, this inequality means that if F_X has a point of discontinuity $x_{j,k} < x_0 < x_{j,k+1}$, then $F_X(x_0+0) - F_X(x_0-0) = b > 0$, i.e. $m(x_0) = b$. Because this discontinuity is between two zeros, we thus find $\lambda_{j,k} + \lambda_{j,k+1} \geq b$ for all j.

On the contrary, if F_X possess certain properties of continuity, $m^j \xrightarrow{d} m$. Moreover $\lambda_{j,k}$'s converges regularly to 0 and distances of successive zeros $x_{j,k}$ converges to 0 (cf [14]). Some of these properties are grouped together in [9] : http://hal.archives-ouvertes.fr/hal-00587108/fr/.

These theorems in particular means that if there is no point x_0 such that $m(\{x_0\}) > 0$, the distribution of zeros and weights is enough regular. Because this is not the case if $m(\{x_0\}) > 0$, it will detect the existence of those discontinuities by a way enough simple.

In particular, there cannot be more than two zeros in an interval of measure zero.

Proposition 1.3.2 *It can not be three successive zeros $x_{j,s} < x_{j,s+1} < x_{j,s+2}$ such that $P\{X \in [x_{j,s}, x_{j,s+2}]\} = 0$ if $\lambda_{j,s+1} > 0$.*

Proof By Stieljes Markov inequality, we know that $\sum_{x_{j,s}<x_{j,k+2}} \lambda_{j,s} \leq F_X(x_{j,k+2}-0)$ and $\sum_{x_{j,s}\leq x_{j,k}} \lambda_{j,s} \geq F_X(x_{j,k}+0)$. Then,

$$0 = F_X(x_{j,k+2}) - F_X(x_{j,k}) = F_X(x_{j,k+2}-0) - F_X(x_{j,k}+0)$$

$$\geq \sum_{x_{j,s}<x_{j,k+2}} \lambda_{j,s} - \sum_{x_{j,s}\leq x_{j,k}} \lambda_{j,s} = \lambda_{j,k+1} > 0 \ . \blacksquare$$

Now, the variance of order j is invariant by translation.

Proposition 1.3.3 *Let $a \in \mathbb{R}$. Let m_a the translated probability : $m_a(B) = P(X + a \in B)$. For each $j \in \Theta$, the (j+1)-th orthonormal polynomial associated at m_a is $\tilde{P}_j(x - a)$. Moreover, let $x'_{j,1}, x'_{j,2}, ..., x'_{j,j}$, the zeros of $\tilde{P}_j(x - a)$, $\lambda'_{j,1}, \lambda'_{j,2}, ..., \lambda'_{j,j}$, be the weights of associated Gauss-Jacobi Quadrature, and σ'^2_j be the variance of order j associated at m_a. Then, $x'_{j,s} = x_{j,s} + a$, $\lambda'_{j,s} = \lambda_{j,s}$ and $\sigma'^2_j = \sigma^2_j$.*

In order to prove this result, it is enough to remark that $\int \tilde{P}_j(x - a)\tilde{P}_k(x - a).m_a(dx) = \int \tilde{P}_j(x)\tilde{P}_k(x).m(dx)$.

The following proposition results from the Gram-Schmidt Process

Proposition 1.3.4 *The real σ_j is the distance in $L^2(\mathbb{R}, m)$ of the polynomial $x \longmapsto x^j$ to the subspace of $L^2(\mathbb{R}, m)$ spanned by the polynomials of degree more little than j-1. Moreover, the minimum of $\int \left[(x-t_1)(x-t_2)...(x-t_j)\right]^2.m(dx)$ when $(t_1, t_2, ..., t_j) \in \mathbb{R}^j$ is reached for $(t_1, t_2, ..., t_j) = (x_{j,1}, x_{j,2}, ..., x_{j,j})$ and is equal to σ^2_j .*

Now recall how to calculate practically the variance of order j.

Proposition 1.3.5 *Let $j \in \Theta$. Then,*

$$\sigma^2_j = M_{2j} - \sum_{s=0}^{j-1} \beta^2_{j,s} \text{ where } \beta_{j,s} = \int x^j P_s(x).mdx\,.$$

Proof We have

$$\tilde{P}_j = x^j - \sum_{s=0}^{j-1} \mathbb{E}\{X^j P_s(X)\}P_s(x)\ .$$

Therefore,

$$\sigma^2_j = \int \Big(x^j - \sum_{s=0}^{j-1} \mathbb{E}\{X^j P_s(X)\}P_s(x)\Big)^2 m(dx)$$

$$= \int x^{2j} m(dx) - 2\int x^j \Big(\sum_{s=0}^{j-1} \mathbb{E}\{X^j P_s(X)\}P_s(x)\Big) m(dx) + \int \Big(\sum_{s=0}^{j-1} \mathbb{E}\{X^j P_s(X)\}P_s(x)\Big)^2 m(dx)$$

$$= \int x^{2j} m(dx) - 2\sum_{s=0}^{j-1} \mathbb{E}\{X^j P_s(X)\}^2 + \sum_{s=0}^{j-1} \mathbb{E}\{X^j P_s(X)\}^2\ . \blacksquare$$

We specify now how the $a_{j,t}$'s are obtained.

Proposition 1.3.6 *For all $j \in \mathbb{N}$, for all $t \in \{1, 2, ..., j\}$,*

$$\tilde{P}_j(x) = \sum_{t=0}^{j} a_{j,t} x^t = \sum_{t=0}^{j} O_{j-t}(X) x^t\ ,$$

when $O_s(X) = \frac{N_r(X)}{D_t(X)}$ where $r - t = s$ and where N_r and D_r mean any polynomial, function of the M_i's which are sum or difference of terms in the form $M_0^{n_0} M_1^{n_1} M_2^{n_2} .....M_q^{n_q}$, $0n_0 + 1n_1 + 2n_2 + ... + qn_q = r$, $n_i \in \mathbb{N}$.

Proof Remark that, by our definition $N_t(X)N_s(X) = N_{s+t}(X)$, $D_t(X)D_s(X) = D_{s+t}(X)$, $N_t(X) + N_t(X) = N_t(X)$ and $D_t(X) + D_t(X) = D_t(X)$. Therefore, $O_t(X)O_s(X) = O_{s+t}(X)$ and $O_t(X) + O_t(X) = O_t(X)$.

Proposition holds when j=0,1,2. Thus, $\tilde{P}_0 \equiv 1$, $\tilde{P}_1(x) = x - O_1(X) = O_0(X)x - O_1(X)$ where $O_1(X) = M_1$, and $O_0(X) = M_0$.

Moreover, $\tilde{P}_2(x) = x^2 - \frac{M_3 - M_1 M_2}{M_2 - M_1^2}(x - M_1) - M_2 = x^2 - O_1(X)(x - M_1) - O_2(X) = O_0(X)x^2 - O_1(X)x - O_2(X)$.

Then, we suppose that $\tilde{P}_j(x) = \sum_{t=0}^{j} O_{j-t}(X)x^t$ for all $j < J$.

Therefore,

$$\tilde{P}_j(x)^2 = \sum_{t,t'=0}^{j} O_{j-t}(X)O_{j-t'}(X)x^t x^{t'} = \sum_{t"=0}^{2j} \sum_{t+t'=t"} O_{j-t}(X)O_{j-t'}(X)x^t x^{t'}$$

$$= \sum_{t"=0}^{2j} \sum_{t+t'=t"} O_{2j-t"}(X)x^{t"} = \sum_{t"=0}^{2j} O_{2j-t"}(X)x^{t"} \ .$$

Therefore,

$$\int \tilde{P}_j(x)^2.m(dx) = \sum_{t"=0}^{2j} O_{2j-t"}(X) \int x^{t"}.m(dx) = \sum_{t"=0}^{2j} O_{2j-t"}(X)O_{t"}(X) = O_{2j}(X) \ .$$

Therefore, $\sigma_j(X)^2 = O_{2j}(X)$.

Then,

$$\frac{\tilde{P}_s(x)}{\sigma_s^2} = \sum_{t=0}^{s} \frac{O_{s-t}(X)}{O_{2s}(X)} x^t = \sum_{t=0}^{s} O_{-s-t}(X)x^t \ .$$

Therefore,

$$\frac{x^J \tilde{P}_s(x)}{\sigma_s^2} = \sum_{t=0}^{s} O_{-s-t}(X)x^{t+J} \ .$$

Because $P_s(x) = \frac{\tilde{P}_s(x)}{\sigma_s}$, then,

$$\frac{\mathbb{E}\{X^J P_s(X)\}}{\sigma_s} = \frac{\mathbb{E}\{X^J \tilde{P}_s(X)\}}{\sigma_s^2} = \sum_{t=0}^{s} O_{-s-t}(X)M_{t+J} = O_{J-s}(X) \ .$$

Then, by the Gram-Schmidt Process and by recurrence,

$$\tilde{P}_J(x) = x^j - \sum_{s=0}^{J-1} \mathbb{E}\{X^J P_s(X)\}P_s(x) = x^j - \sum_{s=0}^{J-1} \frac{\mathbb{E}\{X^j \tilde{P}_s(X)\}}{\sigma_s^2} \tilde{P}_s(x)$$

$$= x^J - \sum_{s=0}^{J-1} O_{J-s}(X)\Big[\sum_{t=0}^{s} O_{s-t}(X)x^t\Big] = x^J - \sum_{s=0}^{J-1}\Big[\sum_{t=0}^{s} O_{J-t}(X)x^t\Big]$$

$$= x^J - \sum_{s=0}^{J-1} O_{J-s}(X)x^s = \sum_{s=0}^{J} O_{J-s}(X)x^s \ . \blacksquare$$

We are now interested in the orthogonal polynomials associated to αX.

Proposition 1.3.7 *We denote by $P_j^\alpha = \sum_{s=0}^{j} c_{j,s}^\alpha x^s$ the orthonormal polynomials associated to αX, $\alpha \in \mathbb{R}_+$. For all $j \in \mathbb{N}$, $j < n_0^m$, for all $s \in \{1, 2, ..., j\}$, we set $c_{j,s}^\alpha = \alpha^s c_{j,s}$. Then,*

$$P_j(x) = \sum_{t=0}^{j} c_{j,t} x^t \ .$$

Proof Let $Q_j(x) = \sum_{s=0}^{j} c_{j,s} x^s$. Then,

$$\delta_{i,j} = \mathbb{E}\big\{P_j^\alpha(\alpha X) P_i^\alpha(\alpha X)\big\}$$

$$= \mathbb{E}\Big\{\Big[\sum_{t=0}^{j} c_{j,t}^\alpha (\alpha X)^t\Big]\Big[\sum_{s=0}^{i} c_{i,s}^\alpha (\alpha X)^s\Big]\Big\}$$

$$= \sum_{t=0}^{j}\sum_{s=0}^{i} c_{j,t}^\alpha c_{i,s}^\alpha \mathbb{E}\big\{(\alpha X)^{t+s}\big\}$$

$$= \sum_{t=0}^{j}\sum_{s=0}^{i} c_{j,t} c_{i,s} \mathbb{E}\{X^{t+s}\}$$

$$= \mathbb{E}\big\{Q_j(X) Q_i(X)\big\} \ .$$

Then, the polynomials $Q_j(x) = \sum_{s=0}^{j} c_{j,s} x^s$ are orthonormal with respect to the measure m. Now orthonormal polynomials such that $c_{j,j} > 0$ are determined by recurrence by the process of Gram Schmidt and are unique (cf Corollary page 9 of [13]). Then, $Q_j = P_j$.

For example, one can prove by recurrence that they are unique. One sets $\tilde{Q}_j(x) = \sum_{s=0}^{j} \frac{c_{j,s}}{c_{j,j}} x^s = \sum_{s=0}^{j-1} \frac{c_{j,s}}{c_{j,j}} x^s + x^j$. By the orthogonality of the Q_j's , $-\sum_{s=0}^{j-1} \frac{c_{j,s}}{c_{j,j}} x^s$ is the orthogonal projection of $x \mapsto x^j$ onto the subspace spanned by $x \mapsto x^s$, $s < j$. This one is unique. Then, by Gram Schmidt Process, $\tilde{Q}_j(x) = \tilde{P}_j(x)$.

Then, $1 = \mathbb{E}\big\{Q_j(X)^2\big\} = \mathbb{E}\big\{[c_{j,j}\tilde{Q}_j(X)]^2\big\} = \mathbb{E}\big\{[c_{j,j}\tilde{P}_j(X)]^2\big\} = c_{j,j}^2 \sigma_j^2$.

Therefore, $c_{j,j}^2 = \frac{1}{\sigma_j^2}$. Therefore, $Q_j(x) = P_j(x)$.

Therefore, $Q_j(x) = \sum_{s=0}^{j} c_{j,s} x^s = P_j(x)$ where $c_{j,s}^\alpha \alpha^s = c_{j,s}$. $\blacksquare$

Exemple j=0 : In this case, $c_{0,0}^\alpha \alpha^0 = c_{0,0} = 1 : Q_0(x) = P_0(x) = 1$.

Exemple j=1 : In order to compare better, it is necessary to use the proposition 1.3.8 below. Then, we remark that

$$\frac{\mathbb{E}\{X\}}{\sigma_1} = \frac{\mathbb{E}\{\alpha X\}}{\sigma_1^\alpha} = c_{1,0}^\alpha = c_{1,0}^\alpha \alpha^0 \overset{def}{=} c_{1,0} = \frac{\mathbb{E}\{X\}}{\sigma_1},$$

$$\frac{1}{\sigma_1} = \frac{\alpha}{\sigma_1^\alpha} = c_{1,1}^\alpha \alpha^1 \overset{def}{=} c_{1,1} = \frac{1}{\sigma_1}$$

Proposition 1.3.8 *Let $\sigma_j(\alpha X)^2$ be the variance of order j associated to αX . Then,*

$$\sigma_j(\alpha X)^2 = \alpha^{2j}\sigma_j^2 \ .$$

Proof We know that $c_{j,j}^2 = \frac{1}{\sigma_j^2}$, $Q_j(x) = P_j(x)$ and $c_{j,s}^\alpha \alpha^s = c_{j,s}$.

Therefore,

$$P_j^\alpha(x) = \sum_{s=0}^{j} c_{j,s}^\alpha x^s = \sum_{s=0}^{j} \frac{c_{j,s}}{\alpha^s} x^s \ .$$

Therefore,

$$\tilde{P}_j^\alpha(x) = \sum_{s=0}^{j} \frac{c_{j,s}^\alpha}{c_{j,j}^\alpha} x^s = \sum_{s=0}^{j} \frac{c_{j,s}\alpha^j}{c_{j,j}\alpha^s} x^s = \sum_{s=0}^{j} c_{j,s}\alpha^{j-s}\sigma_j x^s \ .$$

Therefore,

$$\sigma_j(\alpha X)^2 = \mathbb{E}\big\{\tilde{P}_j^\alpha(\alpha X)^2\big\} = \sum_{s=0}^{j}\sum_{t=0}^{j} c_{j,s}\alpha^{j-s} c_{j,t}\alpha^{j-t}\sigma_j^2 \mathbb{E}\big\{(\alpha X)^{s+t}\big\}$$

$$= \alpha^{2j}\sigma_j^2 \sum_{s=0}^{j}\sum_{t=0}^{j} c_{j,s}c_{j,t}\mathbb{E}\Big\{\frac{(\alpha X)^{s+t}}{\alpha^{t+s}}\Big\}$$

$$= \alpha^{2j}\sigma_j^2 \sum_{s=0}^{j}\sum_{t=0}^{j} c_{j,s}c_{j,t}\mathbb{E}\big\{X^{s+t}\big\} = \alpha^{2j}\sigma_j^2\mathbb{E}\big\{P_j(X)^2\big\} = \alpha^{2j}\sigma_j^2 \ .$$

Therefore,

$$\sigma_j(\alpha X)^2 = \alpha^{2j}\sigma_j^2 \ . \blacksquare$$

Now we are interested by the quadrature of Gauss Jacobi of αX.

Proposition 1.3.9 *Let $\tilde{P}_j^\alpha(x) = \sum_{s=0}^{j} a_{j,s}^\alpha x^s$ be the orthogonal polynomials associated to αX such that $a_{j,j}^\alpha = 1$. Let $x_{j,s}^\alpha$ and $\lambda_{j,s}^\alpha$, s=1,2,....,j, be the zeros and the weigths of the quadrature of Gauss Jacobi associated to αX. Then for all $j \in \mathbb{N}^*$, for all $s \in \{1, 2, ..., j\}$,*

$$x_{j,s}^\alpha = \alpha x_{j,s} \quad \textit{and} \quad \lambda_{j,s}^\alpha = \lambda_{j,s} \ .$$

Proof By proposition, 1.3.8,

$$\sum_{s=0}^{j} a_{j,s}^\alpha x^s = \tilde{P}_j^\alpha(x) = \sigma_j^\alpha P_j^\alpha(x) = \alpha^j\sigma_j P_j^\alpha(x)$$

$$= \alpha^j\sigma_j \sum_{s=0}^{j} c_{j,s}^\alpha x^s = \alpha^j\sigma_j \sum_{s=0}^{j}(c_{j,s}/\alpha^s)x^s = \alpha^j \sum_{s=0}^{j} \frac{a_{j,s}}{\alpha^s} x^s \ .$$

Therefore, $a^{\alpha}_{j,s} = \frac{\alpha^j a_{j,s}}{\alpha^s}$.

Therefore,

$$\sum_{s=0}^{j} a^{\alpha}_{j,s}\alpha^s x^s_{j,t} = \sum_{s=0}^{j} \frac{\alpha^j a_{j,s}\alpha^s x^s_{j,t}}{\alpha^s} = \alpha^j \sum_{s=0}^{j} a_{j,s} x^s_{j,t} = 0 \ .$$

Therefore the zeros of $\tilde{P}^{\alpha}_j$ are the $\alpha x_{j,t}$'s.

In order to study the weigths $\ell^j_s(x)$,it is enough to study $\ell^j_1(x) = \frac{(x-x_{j,2})(x-x_{j,3})...(x-x_{j,j})}{(x_{j,1}-x_{j,2})(x_{j,1}-x_{j,3})...(x_{j,1}-x_{j,j})}$. Indeed one deduces the other $\ell^j_s(x)$'s by permuting the definition of the $x_{j,t}$'s.

Therefore,

$$\lambda_{j,1} = \int \frac{(x-x_{j,2})(x-x_{j,3})...(x-x_{j,j})}{(x_{j,1}-x_{j,2})(x_{j,1}-x_{j,3})...(x_{j,1}-x_{j,j})} m(dx) \ .$$

.

Moreover,

$$\lambda^{\alpha}_{j,1} = \int \frac{(x-x^{\alpha}_{j,2})(x-x^{\alpha}_{j,3})...(x-x^{\alpha}_{j,j})}{(x^{\alpha}_{j,1}-x^{\alpha}_{j,2})(x^{\alpha}_{j,1}-x^{\alpha}_{j,3})...(x^{\alpha}_{j,1}-x^{\alpha}_{j,j})} m_{\alpha}(dx)$$

.

$$= \int \frac{(x-x^{\alpha}_{j,2})(x-x^{\alpha}_{j,3})...(x-x^{\alpha}_{j,j})}{\alpha^{j-1}(x_{j,1}-x_{j,2})(x_{j,1}-x_{j,3})...(x_{j,1}-x_{j,j})} m_{\alpha}(dx)$$

.

$$= \frac{\int \Big[x^{j-1} - (x^{\alpha}_{j,2} + x^{\alpha}_{j,3} + ... + x^{\alpha}_{j,j})x^{j-2} + ... + (-1)^{j-1}x^{\alpha}_{j,2}x^{\alpha}_{j,3}.....x^{\alpha}_{j,j}\Big]}{\alpha^{j-1}(x_{j,1}-x_{j,2})(x_{j,1}-x_{j,3})...(x_{j,1}-x_{j,j})} m_{\alpha}(dx)$$

$$= \frac{\alpha^{j-1}M_{j-1} - (x^{\alpha}_{j,2} + x^{\alpha}_{j,3} + ... + x^{\alpha}_{j,j})\alpha^{j-2}M_{j-2} + ... + (-1)^{j-1}x^{\alpha}_{j,2}x^{\alpha}_{j,3}.....x^{\alpha}_{j,j}\alpha^0 M_0\Big]}{\alpha^{j-1}(x_{j,1}-x_{j,2})(x_{j,1}-x_{j,3})...(x_{j,1}-x_{j,j})}$$

$$= \frac{\alpha^{j-1}M_{j-1} - \alpha^1(x_{j,2} + x_{j,3} + ... + x_{j,j})\alpha^{j-2}M_{j-2} + ... + (-1)^{j-1}\alpha^{j-1}x_{j,2}x_{j,3}.....x_{j,j}\alpha^0 M_0\Big]}{\alpha^{j-1}(x_{j,1}-x_{j,2})(x_{j,1}-x_{j,3})...(x_{j,1}-x_{j,j})}$$

$$= \frac{M_{j-1} - (x_{j,2} + x_{j,3} + ... + x_{j,j})M_{j-2} + ... + (-1)^{j-1}x_{j,2}x_{j,3}.....x_{j,j}M_0\Big]}{(x_{j,1}-x_{j,2})(x_{j,1}-x_{j,3})...(x_{j,1}-x_{j,j})}$$

$$= \lambda_{j,1} \ . \ \blacksquare$$

Chapter 2

Estimation

We will see that one can easily estimate the higher order variances and the Gauss Jacobi quadrature. We can also obtain their asymptotic distributions because we know the asymptotic distributions of empirical orthogonal functions.

At first, we need empirical orthogonal polynomials.

Notations 2.0.1 *Let $h \in \mathbb{N}$ such that $\tilde{P}_h$ exists. Let $\{X_\ell\}_{\ell\in\mathbb{N}}$, $X_\ell \in \mathbb{R}$, be an IID sequence of random variables where $X_0 = X$.*

For all $n \in \mathbb{N}^*$, we denote by m_n the empirical measure associated at the sample $\{X_\ell\}_{\ell=1,2,..,n}$.

Notations 2.0.2 *Let $\{\tilde{P}_j^n\}_{j=0,1,..,h}$ and $\{P_j^n\}_{j=0,1,..,h}$ be the family of orthogonal polynomials associated to the empirical measure m_n such that $\int (P_j^n)^2.dm_n = 1$ if P_j^n exists and $\tilde{P}_j^n = \sum_{s=0}^{j-1} a_{j,s}^n x^s + x^j$ if $\tilde{P}_j^n$ exists. If $\tilde{P}_j^n$ does not exist, we set $\tilde{P}_j^n = 0$ and if P_j^n does not exist, we set $P_j^n = 0$.*

We deduce from theorem 6-4 of [2] that $\tilde{P}_i^n$ is an estimator of $\tilde{P}_i$.

Proposition 2.0.10 *For all $i \leq h$, we set $\tilde{P}_i^n = \tilde{P}_i + \sum_{s=0}^{i-1} \tilde{\alpha}_{i,s}^n P_s$ and $P_i^n = P_i + \sum_{s=0}^{i} \alpha_{i,s}^n P_s$. Then, for all (i,s), $\tilde{\alpha}_{i,s}^n$ and $\alpha_{i,s}^n$ converges almost surely to 0.*

Now, one can define empirical Gauss Jacobi Quadrature and empirical variances of order j.

Notations 2.0.3 *Let $j \in \mathbb{N}$. We denote by $(\sigma_j^n)^2$ the variance of order j of m_n, by $x_{j,1}^n, x_{j,2}^n, ..., x_{j,j}^n$ the zeros of $\tilde{P}_j^n$, by $\lambda_{j,1}^n, \lambda_{j,2}^n, ..., \lambda_{j,j}^n$ the weights of Gauss Jacobi quadrature, if these numbers exists. If not, one defines theses variables by 0.*

Remark that $(\sigma_1^n)^2$ is the classical empirical variances. Moreover $(\sigma_2^n)^2 = M_4^n - \frac{(M_3^n - M_1^n M_2^n)^2}{M_2^n - (M_1^n)^2} - (M_2^n)^2$, where M_j^n is the empirical moment of order j.

Clearly, these estimators converges almost surely if X_ℓ is IID.

Proposition 2.0.11 *Under the previous assumptions, $\sigma_j^n \overset{a.s.}{\to} \sigma_j$. Moreover, for all j=1,2,..,h, for all s=1,2,..,j, $x_{j,s}^n \overset{a.s.}{\to} x_{j,s}$ and $\lambda_{j,s}^n \overset{a.s.}{\to} \lambda_{j,s}$, respectively.*

Proof For example, let us write $\tilde{P}_j(x) = \sum_{t=0}^{j} a_{j,t} x^t$ and $\tilde{P}_j^n(x) = \sum_{t=0}^{j} a_{j,t}^n x^t$. Because $\tilde{\alpha}_{j,s}^n \overset{a.s.}{\to} \tilde{0}$ and $\alpha_{j,s}^n \overset{a.s.}{\to} 0$, then, $a_{j,s}^n \overset{a.s.}{\to} a_{j,s}$.

Now, by theorem page 24 of [12], we can write the following property. $g(a_{j,0}^n, a_{j,1}^n, ..., a_{j,j}^n) \overset{a.s.}{\to} g(a_{j,0}, a_{j,1}, ..., a_{j,j})$ if g is continous with P-probability 1. Now, for example, the zero $x_{j,s}^n$ is written in a form $x_{j,s}^n = g_0(a_{j,0}^n, a_{j,1}^n, ..., a_{j,j}^n)$. Then, $x_{j,s}^n \overset{a.s.}{\to} x_{j,s}$. ■.

Now, by the theorem 6-9 of [2], we have the following theorem.

Theorem 3 *We suppose that* $\mathbb{E}\{X^{4h}\} < \infty$. *Then,* $\alpha^n_{i,s} = -\int P_i P_s.dm_n + o_p\big(n^{-1/2}\big)$ *if* $s < i$ *and* $\alpha^n_{i,i} = \frac{1-\int P_i P_i.dm_n}{2} + o_p\big(n^{-1/2}\big)$ *where* O_p *and* o_p *are the sthocastics "O" and "o"* [1].

Moreover, $\tilde{\alpha}^n_{i,s} = -\int \tilde{P}_i P_s.dm_n + o_p\big(n^{-1/2}\big)$ *if* $s < i$.

Then, one can generalize the result about the asymptotical distribution of empirical classical variance.

Theorem 4 *Under the previous assumptions,* $\sqrt{n}\big[(\sigma^n_j)^2 - \sigma^2_j\big]$ *has asymptotically a normal distribution with mean 0 and variance* $\mathbb{E}\big\{\tilde{P}_j(X)^4\big\} - \sigma^4_j$.

Proof : It is easy to prove that $\sqrt{n}\Big(\int \tilde{P}^n_j \tilde{P}^n_j.dm_n - \int \tilde{P}_j \tilde{P}_j.dm\Big)$ has asymptotically the same distribution as $\sqrt{n}\Big(\int \tilde{P}_j \tilde{P}_j.dm_n - \sigma^2_j\Big) = \sum_{\ell=1}^n \frac{\tilde{P}_j(X_\ell)\tilde{P}_j(X_\ell)-\sigma^2_j}{\sqrt{n}}$. Then, it is enough to use the Central Limit Theorem. ■

We obtain now the asymptotical distribution of the estimators of Gauss Jacobi Quadrature.

Theorem 5 *For all s=1,2,...,j, we set* $x^n_{j,s} = x_{j,s} + \eta_s$. *Then,* $\sqrt{n}\{\eta_s\}_{s=1,..,j}$ *has asymptotically a normal distribution with mean 0 and covariance matrix* $\{G_{s,t}\}_{(s,t)\in\{1,2,...,j\}^2}$ *where*

$$G_{s,t} \;=\; \mathbb{E}\left\{\frac{\Big[\sum_{v=0}^{j-1} P_v(x_{j,s})P_v(X)P_j(X)\Big]\Big[\sum_{v=0}^{j-1} P_v(x_{j,t})P_v(X)P_j(X)\Big]}{\tilde{P}'_j(x_{j,s})\tilde{P}'_j(x_{j,t})}\right\} .$$

This theorem is deduced from the following lemma by using the CLT.

Lemma 2.0.1 *For all s = 1,2,...,j,*

$$\eta_s = \frac{1}{\tilde{P}'_j(x_{j,s})}\int \tilde{P}_j(t)\Big(\sum_{v=0}^{j-1} P_v(x_{j,s})P_v(t)\Big).m_n(dt) + o_p(n^{-1/2}) .$$

Proof We prove this lemma for s=1. First, one proves that $\sqrt{n}\ \eta_1$ is asymptotically normal. We know that

$$\tilde{P}^n_j \;=\; \tilde{P}_j + \sum_{v=0}^{j-1} \tilde{\alpha}^n_{j,v} P_v .$$

Therefore, because $\tilde{P}^n_j(x^n_{j,1}) = 0$,

$$\tilde{P}_j(x^n_{j,1}) \;=\; -\sum_{v=0}^{j-1} \tilde{\alpha}^n_{j,v} P_v(x^n_{j,1}) .$$

Therefore, $\sqrt{n}\tilde{P}_j(x^n_{j,1}) \;=\; \sqrt{n}(x^n_{j,1} - x_{j,1})(x^n_{j,1} - x_{j,2})....(x^n_{j,1} - x_{j,j})$ is asymptotically normal. Moreover $(x^n_{j,1} - x_{j,2})^{-1}....(x^n_{j,1} - x_{j,j})^{-1}$ converges almost surely to $(x_{j,1} - x_{j,2})^{-1}....(x_{j,1} - x_{j,j})^{-1}$. Therefore, by the theorem of page 19 of [12] $\sqrt{n}(x^n_{j,1} - x_{j,1})$ is asymptotically normal, i.e. $\sqrt{n}\ \eta_1$ is asymptotically normal.

By the same way, one proves that $\sqrt{n}\ \eta_s$ is asymptotically normal for s=2,3,...,j.

[1] A sequence of random variable X_n is bounded in probability, if, for every $\epsilon > 0$, there exists M_ϵ and N_ϵ such that $P\{|X_n| \le M_\epsilon\} \ge 1 - \epsilon$ for all $n \ge N_\epsilon$. Then, one writes $X_n = O_P(1)$. Moreover, we write $X_n = O_P(Z_n)$ for two sequences of random variable X_n and Z_n, if $X_n/Z_n = O_P(1)$ and $X_n = o_P(Z_n)$ if $X_n/Z_n \xrightarrow{p} 0$.

Now one can prove the lemma. Indeed, one can write the following equalities :

$$\tilde{P}_j^n(x) = (x - x_{j,1} - \eta_1)(x - x_{j,2} - \eta_2).....(x - x_{j,j} - \eta_j)$$

$$\begin{aligned}
= & \quad (x - x_{j,1})(x - x_{j,2}).....(x - x_{j,j}) \\
- & \quad \eta_1(x - x_{j,2})(x - x_{j,3}).....(x - x_{j,j}) \\
- & \quad \eta_2(x - x_{j,1})(x - x_{j,3}).....(x - x_{j,j}) \\
& \quad ..... \\
- & \quad \eta_j(x - x_{j,1})(x - x_{j,2}).....(x - x_{j,j-1}) + o_p(n^{-1/2}) .
\end{aligned}$$

Therefore,

$$\tilde{P}_j^n(x_{j,1}) = -\eta_1(x_{j,1} - x_{j,2})(x_{j,1} - x_{j,3}).....(x_{j,1} - x_{j,j}) + o_p(n^{-1/2}) = -\eta_1 \tilde{P}_j'(x_{j,1}) + o_p(n^{-1/2}) .$$

Now, by theorem 3,

$$\begin{aligned}
\tilde{P}_j^n(x_{j,1}) &= \tilde{P}_j(x_{j,1}) + \sum_{v=0}^{j-1} \tilde{\alpha}_{j,v}^n P_v(x_{j,1}) \\
&= \sum_{v=0}^{j-1} \tilde{\alpha}_{j,v}^n P_v(x_{j,1}) \\
&= \sum_{v=0}^{j-1} \Big(- \int \tilde{P}_j P_v dm_n \Big) P_v(x_{j,1}) + o_p(n^{-1/2}) .
\end{aligned}$$

We deduce the lemma. ■

Theorem 6 *Let $j \in \mathbb{N}$ such that $\mathbb{E}\{X^{4j}\} < \infty$. For all s=1,2,...,j, and for all u=1,2,..,j, we set $L_s(x) = \frac{\tilde{P}_j(x)}{x - x_{j,s}}$ and $h_u^s(x) = \frac{\tilde{P}_j(x)}{(x - x_{j,s})(x - x_{j,u})}$. We define D_u^s by*

$$\begin{aligned}
D_u^s &= \sum_{r=1, r \neq s}^{j} \left(\frac{h_r^s(x_{j,s}) \mathbb{E}\{L_s(X)\}}{L_s(x_{j,s})} - \mathbb{E}\{h_r^s(X)\} \right) \left(\frac{P_u(x_{j,r})}{L_s(x_{j,s}) \tilde{P}_j'(x_{j,r})} \right) \\
& - \frac{P_u(x_{j,s}) \mathbb{E}\{L_s(X)\}}{L_s(x_{j,s})^3} \sum_{r=1, s \neq r}^{j} h_r^s(x_{j,s}) .
\end{aligned}$$

Then, $\sqrt{n}\{\lambda_{j,s}^n - \lambda_{j,s}\}_{s=1,..,j}$ has asymptotically the normal distribution with mean 0 and covariance matrix $\{O_{s,t}\}_{1 \leq s,t \leq j}$ where

$$O_{s,t} = \mathbb{E}\left\{ \Big[\ell_s(X) + \sum_{u=0}^{j-1} D_u^s P_u(X) \tilde{P}_j(X) \Big] \Big[\ell_t(X) + \sum_{u=0}^{j-1} D_u^t P_u(X) \tilde{P}_j(X) \Big] \right\} - \lambda_{j,s} \lambda_{j,t} .$$

This theorem is deduced from the following lemma by using CLT.

Lemma 2.0.2 *For all $s \in \{1, 2, .., j\}$,*

$$\lambda_{j,s}^n = \int \ell_s(t).m_n(dt) + \int \Big(\sum_{u=0}^{j-1} D_u^s P_u(t) \tilde{P}_j(t) \Big).m_n(dt) + o_p(n^{-1/2}) .$$

Proof We prove this lemma for s=1. We simplify $L^1(t)$ in L(t) and h_r^1 in h_r : $L(t) = (t-x_{j,2})(t-x_{j,3})....(t-x_{j,j})$, $h_2(t) = (t-x_{j,3})(t-x_{j,4})....(t-x_{j,j})$, $h_3(t) = (t-x_{j,2})(t-x_{j,4})....(t-x_{j,j})$ etc. Moreover, we set $L^n(t) = (t-x_{j,2}^n)(t-x_{j,3}^n)....(t-x_{j,j}^n)$.

We know that

$$\begin{aligned}
\lambda_{j,1} &= \int \ell_1(x).m(dx) \\
&= \frac{\int (t-x_{j,2})(t-x_{j,3})...(t-x_{j,j}).m(dt)}{(x_{j,1}-x_{j,2})(x_{j,1}-x_{j,3})...(x_{j,1}-x_{j,j})} \\
&= \frac{\int L(t).m(dt)}{L(x_{j,1})}
\end{aligned}$$

and that

$$\begin{aligned}
\lambda_{j,1}^n &= \int \ell_1^n(x).m_n(dx) \\
&= \frac{\int (t-x_{j,2}^n)(t-x_{j,3}^n)...(t-x_{j,j}^n).m_n(dt)}{(x_{j,1}^n-x_{j,2}^n)(x_{j,1}^n-x_{j,3}^n)...(x_{j,1}^n-x_{j,j}^n)} \\
&= \frac{\int L^n(t).m_n(dt)}{L^n(x_{j,1}^n)} \quad .
\end{aligned}$$

By proposition 2.0.11, we deduce easily that $\int L^n(t).m_n(dt)$ and $L^n(x_{j,1}^n)$ converge almost surely to $\int L(t).m(dt)$ and $L(x_{j,1})$, respectively. Therefore,

$$\lambda_{j,1}^n - \lambda_{j,1} = \frac{\int L^n(t).m_n(dt)}{L^n(x_{j,1}^n)} - \frac{\int L(t).m(dt)}{L(x_{j,1})}$$

$$= \frac{\int L^n(t).m_n(dt) - \int L(t).m(dt)}{L^n(x_{j,1}^n)} + \frac{\int L(t).m(dt)}{L^n(x_{j,1}^n)} - \frac{\int L(t).m(dt)}{L(x_{j,1})}$$

$$= \frac{\int L^n(t).m_n(dt) - \int L(t).m(dt)}{L(x_{j,1})} - \frac{\Big[\int L(t).m(dt)\Big]\Big[L^n(x_{j,1}^n) - L(x_{j,1})\Big]}{L(x_{j,1})^2} + o_p(n^{-1/2}) ,$$

if $\sqrt{n}\big[\int L^n(t).m_n(dt) - \int L(t).m(dt)\big]$ and $\sqrt{n}\big[L^n(x_{j,1}^n) - L(x_{j,1})\big]$ are asymptotically normal. Then, we prove this result now. Indeed,

$$\int L^n(t).m_n(dt) - \int L(t).m(dt)$$

$$= \int (t-x_{j,2}^n)(t-x_{j,3}^n)....(t-x_{j,j}^n).m_n(dt) - \int (t-x_{j,2})(t-x_{j,3})....(t-x_{j,j}).m(dt)$$

$$\begin{aligned}
&= \int (t-x_{j,2})(t-x_{j,3})....(t-x_{j,j}).[m_n-m](dt) \\
&- \eta_2 \int (t-x_{j,3})(t-x_{j,4})....(t-x_{j,j}).m_n(dt)
\end{aligned}$$

$$
\begin{aligned}
&- \quad \eta_3 \int (t - x_{j,2})(t - x_{j,4}) .... (t - x_{j,j}).m_n(dt) \\
&- \quad ..... \\
&- \quad \eta_j \int (t - x_{j,2})(t - x_{j,3}) .... (t - x_{j,j-1}).m_n(dt) \ + \ o_p(n^{-1/2})
\end{aligned}
$$

$$
= \int L(t).[m_n - m](dt) \ - \ \eta_2 \mathbb{E}\{h_2(X)\} - \ \eta_3 \mathbb{E}\{h_3(X)\} \ - \ ..... - \ \eta_j \mathbb{E}\{h_j(X)\} + \ o_p(n^{-1/2}) \quad .
$$

By the same way,

$$
\begin{aligned}
& L^n(x_1^n) \ - \ L(x_{j,1}) \\
= \ & (x_{j,1}^n - x_{j,2}^n)(x_{j,1}^n - x_{j,3}^n) ... (x_{j,1}^n - x_{j,j}^n) \ - \ (x_{j,1} - x_{j,2})(x_{j,1} - x_{j,3}) ... (x_{j,1} - x_{j,j}) \\
= \ & (x_{j,1}^n - x_{j,2})(x_{j,1}^n - x_{j,3}) ... (x_{j,1}^n - x_{j,j}) \ - \ (x_{j,1} - x_{j,2})(x_{j,1} - x_{j,3}) ... (x_{j,1} - x_{j,j}) \\
- \ & \eta_2 (x_{j,1}^n - x_{j,3})(x_{j,1}^n - x_{j,4}) ... (x_{j,1}^n - x_{j,j}) \\
- \ & \eta_3 (x_{j,1}^n - x_{j,2})(x_{j,1}^n - x_{j,4}) ... (x_{j,1}^n - x_{j,j}) \\
- \ & ..... \\
- \ & \eta_j (x_{j,1}^n - x_{j,2})(x_{j,1}^n - x_{j,3}) ... (x_{j,1}^n - x_{j,j-1}) + \ o_p(n^{-1/2})
\end{aligned}
$$

$$
\begin{aligned}
= \ & (x_{j,1} - x_{j,2})(x_{j,1} - x_{j,3}) ... (x_{j,1} - x_{j,j}) \ - \ (x_{j,1} - x_{j,2})(x_{j,1} - x_{j,3}) ... (x_{j,1} - x_{j,j}) \\
+ \ & \eta_1 (x_{j,1} - x_{j,3})(x_{j,1} - x_{j,4}) ... (x_{j,1} - x_{j,j}) \\
+ \ & \eta_1 (x_{j,1} - x_{j,2})(x_{j,1} - x_{j,4}) ... (x_{j,1} - x_{j,j}) \\
+ \ & ..... \\
+ \ & \eta_1 (x_{j,1} - x_{j,2})(x_{j,1} - x_{j,3}) ... (x_{j,1} - x_{j,j-1}) \\
\\
- \ & \eta_2 (x_{j,1} - x_{j,3})(x_{j,1} - x_{j,4}) ... (x_{j,1} - x_{j,j}) \\
- \ & \eta_3 (x_{j,1} - x_{j,2})(x_{j,1} - x_{j,4}) ... (x_{j,1} - x_{j,j}) \\
- \ & ..... \\
- \ & \eta_j (x_{j,1} - x_{j,2})(x_{j,1} - x_{j,3}) ... (x_{j,1} - x_{j,j-1}) + \ o_p(n^{-1/2}) \\
\\
= \ & \eta_1 \big[h_2(x_{j,1}) + ... + h_j(x_{j,1}) \big] \ - \ \eta_2 h_2(x_{j,1}) \ - .... - \ \eta_j h_j(x_{j,1}) + \ o_p(n^{-1/2}) \ .
\end{aligned}
$$

Therefore,

$$
\lambda_{j,1}^n - \lambda_{j,1}
$$

$$
\begin{aligned}
= \ & \frac{\int L(t).[m_n - m](dt)}{L(x_{j,1})} \\
- \ & \eta_2 \frac{\mathbb{E}\{h_2(X)\}}{L(x_{j,1})} \ - \eta_3 \frac{\mathbb{E}\{h_3(X)\}}{L(x_{j,1})} \ - ..... \ - \eta_j \frac{\mathbb{E}\{h_j(X)\}}{L(x_{j,1})} \\
- \ & \eta_1 \frac{\mathbb{E}\{L(X)\}}{L(x_{j,1})^2} \big[h_2(x_{j,1}) + h_3(x_{j,1}) + ... + h_j(x_{j,1}) \big]
\end{aligned}
$$

$$+ \quad \eta_2 h_2(x_{j,1})\frac{\mathbb{E}\{L(X)\}}{L(x_{j,1})^2} + ... + \eta_j h_j(x_{j,1})\frac{\mathbb{E}\{L(X)\}}{L(x_{j,1})^2} + o_p(n^{-1/2})$$

$$\begin{aligned}
&= \frac{\int (t-x_{j,2})(t-x_{j,3})...(t-x_{j,j}).[m_n-m](dt)}{(x_{j,1}-x_{j,2})(x_{j,1}-x_{j,3})...(x_{j,1}-x_{j,j})} \\
&- \eta_1 \frac{\mathbb{E}\{L(X)\}}{L(x_{j,1})^2}\big[h_2(x_{j,1})+h_3(x_{j,1})+....+h_j(x_{j,1})\big] \\
&+ \eta_2 \frac{1}{L(x_{j,1})}\left[\frac{h_2(x_{j,1})\mathbb{E}\{L(X)\}}{L(x_{j,1})} - \mathbb{E}\{h_2(X)\}\right] \\
&+ \eta_3 \frac{1}{L(x_{j,1})}\left[\frac{h_3(x_{j,1})\mathbb{E}\{L(X)\}}{L(x_{j,1})} - \mathbb{E}\{h_3(X)\}\right] \\
&+ ..... \\
&+ \eta_j \frac{1}{L(x_{j,1})}\left[\frac{h_j(x_{j,1})\mathbb{E}\{L(X)\}}{L(x_{j,1})} - \mathbb{E}\{h_j(X)\}\right] + o_p(n^{-1/2})
\end{aligned}$$

$$\begin{aligned}
&= \int \ell_1(x).[m_n-m](dx) \\
&- \left(\int \tilde{P}_j(t)\Big[\sum_{u=0}^{j-1} P_u(x_{j,1})P_u(t)\Big].m_n(dt)\right)\frac{\mathbb{E}\{L(X)\}}{\tilde{P}_j'(x_{j,1})L(x_{j,1})^2}\big[h_2(x_{j,1})+....+h_j(x_{j,1})\big] \\
&+ \left(\int \tilde{P}_j(t)\Big[\sum_{u=0}^{j-1} P_u(x_{j,2})P_u(t)\Big].m_n(dt)\right)\frac{1}{L(x_{j,1})\tilde{P}_j'(x_{j,2})}\left[\frac{h_2(x_{j,1})\mathbb{E}\{L(X)\}}{L(x_{j,1})} - \mathbb{E}\{h_2(X)\}\right] \\
&+ \left(\int \tilde{P}_j(t)\Big[\sum_{u=0}^{j-1} P_u(x_{j,3})P_u(t)\Big].m_n(dt)\right)\frac{1}{L(x_{j,1})\tilde{P}_j'(x_{j,3})}\left[\frac{h_3(x_{j,1})\mathbb{E}\{L(X)\}}{L(x_{j,1})} - \mathbb{E}\{h_3(X)\}\right] \\
&+ ..... \\
&+ \left(\int \tilde{P}_j(t)\Big[\sum_{u=0}^{j-1} P_u(x_{j,j})P_u(t)\Big].m_n(dt)\right)\frac{1}{L(x_{j,1})\tilde{P}_j'(x_{j,j})}\left[\frac{h_j(x_{j,1})\mathbb{E}\{L(X)\}}{L(x_{j,1})} - \mathbb{E}\{h_j(X)\}\right] \\
&+ o_p(n^{-1/2})
\end{aligned}$$

$$\begin{aligned}
&= \frac{\int L(t).[m_n-m](dt)}{L(x_{j,1})} \\
&- \left(\int \tilde{P}_j(t)\Big[\sum_{u=0}^{j-1} P_u(x_{j,1})P_u(t)\Big].m_n(dt)\right)\frac{\mathbb{E}\{L(X)\}}{L(x_{j,1})^3}\big[h_2(x_{j,1})+....+h_j(x_{j,1})\big] \\
&+ \int \Big\{\sum_{u=0}^{j-1}\Big[\sum_{r=2}^{j}\Big(\frac{h_r(x_{j,1})\mathbb{E}\{L(X)\}}{L(x_{j,1})} - \mathbb{E}\{h_r(X)\}\Big)\frac{P_u(x_{j,r})}{L(x_{j,1})\tilde{P}_j'(x_{j,r})}\Big]P_u(t)\tilde{P}_j(t)\Big\}.m_n(dt) \\
&+ o_p(n^{-1/2})\ .
\end{aligned}$$

We deduce the lemma. ■

Chapter 3

Detection of points of concentration

3.1 Introduction

The theorems about orthogonal polynomials of [14] show that in the case of sufficiently smooth density, the distance of successive zeros $x_{j,s}$ converge to 0 and the same is true for weights $\lambda_{j,s}$ (cf [9]. cf also section 1.3). We can ask ourselves what happens when X has a density of this type except at one point x_0 where it has a nonzero measure : $P\{X = x_0\} = b > 0$. We can easily have a first answer thanks to Stieltjes Markov's inequality.

But, for further details, it seems difficult to get them quickly by mathematical theorems. Then the simplest is to make simulations : in all simulations which we have made, we found that the properties of Section 1.3 remain true except near points of nonzero measure. That is to say that, generally, $x_{j,s+1} - x_{j,s} \to 0$ and $\lambda_{j,s} \to 0$ as $j \to \infty$.

But it is quite clear that it will be different for the points closest to x_0 such that $P\{X = x_0\} = b > 0$. Indeed, suppose that two consecutive zeros check $x_{j,k} < x_0 < x_{j,k+1}$ with $P\{X = x_0\} = b > 0$. Then, by Stieljes Markov inequality,

$$F(x_0 + 0) - F(x_0 - 0) \leq F(x_{j,k+1} + 0) - F(x_{j,k} - 0)$$

$$\leq \sum_{x_{j,s} \leq x_{j,k+1}} \lambda_{j,s} - \sum_{x_{j,s} < x_{j,k}} \lambda_{j,s} = \lambda_{j,k} + \lambda_{j,k+1} \ .$$

Therefore, $\lambda_{j,k} + \lambda_{j,k+1} \geq b$. This means that there will always weights which are not too small even if j is large whereas other weights $\lambda_{j,k'}$ converge to 0.

In fact, in some simulations which we have made, we found enough frequently some weights $\lambda_{j,k_j} \to b$ as $j \to \infty$. Moreover, $x_{j,k_j+1} \approx x_0$ ou $x_{j,k_j} \approx x_0$ and $x_{j,k+1} - x_{j,k} \to 0$ as $j \to \infty$.

So it seems that in many cases, in order to find the non-zero measure in a point and its weight, it suffices to find the zeros x_{j,k_j} close to a x_0 and weights $\lambda_{j,k_j} \approx b$.

This is what will confirm the following examples (For other examples cf [9]).

3.2 Example 1

In this example we chose a probability mixing a continuous density and a point $x_0 = 4$ such that $P\{X = x_0\} > 0$. The continuous density is that of a Gaussian mixture centered in -3, -1, with standard deviation 1, 2. The weights of the two Gaussian density and of the Dirac measure (in -3, -1, 4) are the weights 2/9, 4/9, 1/3.

We have calculated the $x^n_{j,t}$'s and the $\lambda^n_{j,t}$'s empirically, i.e. we used a sample of this mixture. Then the function histogram of Matlab for the empirical density is represented in Figure 3.1 [1] .

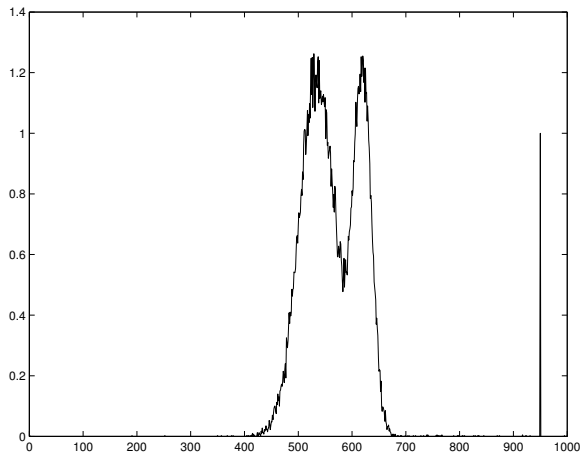

Figure 3.1: Gaussian mixture in -4,-3 and Dirac measure in 4

Then, the points and weights of Gauss Jacobi quadrature allows to detect the concentration at the point $x_0 = 4$. Indeed, the parameters of the Gauss Jacobi quadrature are given by the following tables.

	1	2	3	4	5	6	7	8
x_{1s}	-0.2274	0	0	0	0	0	0	0
x_{2s}	3.4109	-3.2730	0	0	0	0	0	0
x_{3s}	3.9011	-4.8453	-1.7292	0	0	0	0	0
x_{4s}	-6.3918	3.9776	-3.3357	-0.4282	0	0	0	0
x_{5s}	-7.7252	3.9960	-4.5646	-2.2141	0.6129	0	0	0
x_{6s}	-8.7553	-5.6417	4.0051	-3.2739	1.7758	-0.9448	0	0
x_{7s}	-9.5629	-6.5870	4.0981	3.8168	-4.1424	-2.0990	0.3619	0
x_{8s}	-10.2119	-7.4893	5.7010	3.9945	-5.0952	-3.1244	1.1844	-1.1042
x_{9s}	-10.6159	-8.2123	6.2676	-5.9674	3.9995	-3.9839	1.8163	-2.2345
x_{10s}	-10.8087	-8.6956	6.4217	-6.6345	-4.7074	4.0022	-3.0406	2.5305
x_{11s}	-10.8938	-9.0235	6.4882	-7.1783	-5.3689	4.0106	3.6130	-3.7158
x_{12s}	-10.9303	-9.2567	6.5310	-7.6459	-5.9684	4.6708	3.9965	-4.3445

	9	10	11	12
x_{9s}	-0.2326	0	0	0
x_{10s}	-1.3096	0.6104	0	0
x_{11s}	-2.1612	1.4062	-0.4045	0
x_{12s}	1.9751	-2.8888	-1.3550	0.3182

[1] Note that in this figure, it is simultaneously represented the curve of the density of mixture Gaussian and the non-zero measure in $x_0 = 4$. In order that graphs are consistent, it would be required that the Dirac measure is infinite. Normally, it's impossible.

	1	2	3	4	5	6	7	8
$\lambda_{1,s}$	1.0000	0	0	0	0	0	0	0
$\lambda_{2,s}$	0.4557	0.5443	0	0	0	0	0	0
$\lambda_{3,s}$	0.3658	0.1791	0.4551	0	0	0	0	0
$\lambda_{4,s}$	0.0317	0.3454	0.3893	0.2336	0	0	0	0
$\lambda_{6,s}$	0.0052	0.3383	0.1696	0.3761	0.1108	0	0	0
$\lambda_{6,s}$	0.0011	0.0557	0.3333	0.3348	0.0431	0.2320	0	0
$\lambda_{7,s}$	0.0003	0.0176	0.1909	0.1529	0.2029	0.3218	0.1137	0
$\lambda_{8,s}$	0.0001	0.0050	0.0002	0.3375	0.0821	0.3057	0.0579	0.2116
$\lambda_{9,s}$	0.0000	0.0016	0.0000	0.0296	0.3355	0.1943	0.0298	0.2817
$\lambda_{10,s}$	0.0000	0.0007	0.0000	0.0123	0.1010	0.3337	0.2708	0.0138
$\lambda_{11,s}$	0.0000	0.0004	0.0000	0.0055	0.0484	0.3210	0.0175	0.2019
$\lambda_{12,s}$	0.0000	0.0002	0.0000	0.0026	0.0232	0.0008	0.3351	0.1223

	9	10	11	12
$\lambda_{9,s}$	0.1273	0	0	0
$\lambda_{10,s}$	0.1954	0.0723	0	0
$\lambda_{11,s}$	0.2453	0.0383	0.1217	0
$\lambda_{12,s}$	0.0208	0.2452	0.1740	0.0756

This shows that the measure nonzero in $x_0 = 4$ with a weight 1/3 implies zeros close to 4 and weights close to 0,333 while the other zeros and weights are much more varied.

Note that some weights are very close to 0. This is due to properties of pure Gaussian mixtures (without Dirac measure): cf section 3.1.3 [9]

In order to better appreciate these results, in the following table, we shall group together the zeros close to 4 and the associated weight by using the following convention.

Convention 3.2.1 *In the tables giving the two zeros around points x_s with measures $m(\{x_s\}) > 0$, we are only interested by the zeros with significant weight. When there is an alone significant weight we put a 0 for the following weight and we did the same thing for the zero.*

There are then for the zeros close to x=4 and the associated weights the following table.

j	$x_{j,k}$	$x_{j,k+1}$	$\lambda_{j,k}$	$\lambda_{j,k+1}$
3	3.9011	0	0.3658	0
4	3.9776	0	0.3454	0
5	3.9960	0	0.3383	0
6	4.0051	0	0.3333	0
7	4.0981	3.8168	0.1909	0.1529
8	3.9945	0	0.3375	0
9	3.9995	0	0.3355	0
10	4.0022	0	0.3337	0
11	4.0106	0	0.3210	0
12	3.9965	0	0.3351	0

By Stieljes Markov's inequality, we know that the weight at x = 4, checks $\lambda_{j,k} + \lambda_{j,k+1} \geq P\{X = 4\}$ when $x_{j,k} < 4 < x_{j,k+1}$. So we see that there is a weight of about 0,333 concentrated around 4: indeed, $P\{X \in [4-a, 4+a]\} \leq \lambda_{j,k} + \lambda_{j,k+1}$ where a is small enough .

Remark 3.2.2 *One can find practical problems for calculating the elements of the Gauss Jacobi quadrature. Indeed, it is known that the computation of orthogonal polynomials P_j can be difficult when j increases a little. This is because the moments become very large or very small depending on the case. The accuracy of calculations becomes delicate and one can sometimes find values enough distant of real values.*

3.3 Example 2

We choose uniform distributions on two intervals and a point of nonzero measure located inside of their supports. Then we take an uniform mixture of the intervals [-3.0], [0,0.750] and a non-zero measure in -1 with weights, respectively 0.3, 0.5, 0.2.

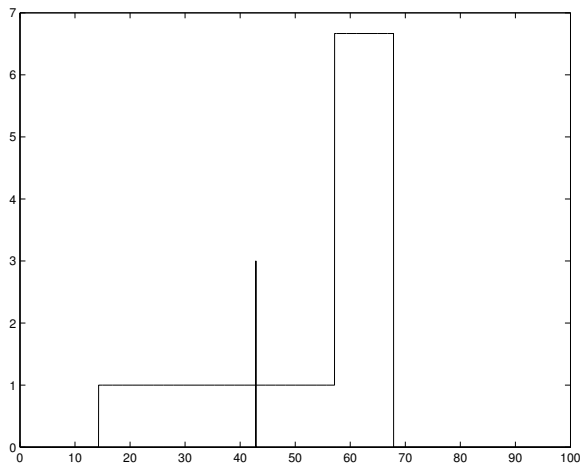

Figure 3.2: One point with a nonzero measure

zeros close to -1 and the associated weights are given in the following table.

j	$x_{j,k}$	$x_{j,k+1}$	$\lambda_{j,k}$	$\lambda_{j,k+1}$
5	-1.0055	0	0.3157	0
6	-1.3163	-0.7185	0.1797	0.2035
7	-1.0081	0	0.2852	0
8	-1.2715	-0.8540	0.1320	0.2036
9	-1.0145	0	0.2650	0
10	-1.2385	-0.9095	0.2067	0.0755
11	-1.0156	0	0.2521	0
12	-1.2521	-0.9544	0.0695	0.2228
13	-1.0389	0	0.2234	0
14	-1.0308	0	0.2310	0
15	-0.9705	0	0.2273	0
16	-1.0026	-0	0.2422	0

The points of concentration and weights appear clearly: they are the zeros close to x=0 and the weights associated.

Chapter 4

Application : mixtures

4.1 Some properties

At first we introduce the type of mixtures that we will study.

Notations 4.1.1 *Let Y be a real random variable. We assume that there exists $p \in \mathbb{N}$ such that $Y = \sum_{r=1}^{p} \delta_r Z_r$, where $Z_r - z_r = \sigma_r \epsilon_r \sim N(0, \sigma_r^2)$, $z_r \in \mathbb{R}$, where the δ_r are random variables such that $\delta_r = 0$ or 1 , $\sum_{r=1}^{p} \delta_r = 1$, and where $(\delta_1, \delta_2, ..., \delta_p), \epsilon_1, \epsilon_2, ..., \epsilon_p$ are independent.*

We assume that the z_r's , $r = 1, 2, .., p$, are distincts. We set $P(\delta_r = 1) = q_r$ for $r = 1, 2, ..., p$.

One want to estimate the z_r 's and q_r's by using Gauss Jacobi Quadrature. Unfortunately, the Gauss-Jacobi quadrature of Y does not give the z_r 's and q_r's directly. Indeed, let us denote by $y_{p,s}$ and $\lambda'_{p,s}$, the zeros and the weights of orthogonal polynomials associated to Y. Then, generally, $y_{p,s} \neq z_s$ and $\lambda'_{p,s} \neq q_s$.

It is true only when $\sigma_p(Y) = 0$. In other cases, we have only approximations which are sharp all the more as $\sigma_p(Y)$ is small.

Proposition 4.1.1 *We denote by $\{\tilde{Q}_j\}$ the familly of orthogonal polynomials associated with $\sum_{r=1}^{p} \delta_r z_r$: $\tilde{Q}_j(z) = \sum_{t=0}^{j} b_{j,t} z^t$ where $b_{j,j} = 1$. Then, the z_r's are the zeros $z_{p,r}$ of $\tilde{Q}_p$. Moreover, let $\ell_r^S(z) = \frac{\tilde{Q}_p(z)}{(z-z_r)\tilde{Q}'_p(z_r)}$. Then, $q_r = \int \ell_r^S(z).m_S(dz)$ for $r = 1, 2...., p$ where m_S is the distribution of $S = \sum_{r=1}^{p} \delta_r z_r$.*

Therefore, we have to use the Gauss Jacobi Quadrature not of Y but of $S = \sum_{r=1}^{p} \delta_r z_r$. Therefore, we need to know moments of $\sum_{r=1}^{p} \delta_r z_r$. In this aim, one has the idea to use "negative variance" of Gaussian random variables (cf Section V-2, page IV-28, [6]) .

Indeed, recall that this naming is based on the following property (cf lemma 1.1, (viii) [10] or [9]).

Proposition 4.1.2 *Soit $G \sim N(0, \sigma^2)$, $U \sim N(0, 1)$ and $\epsilon \sim N(0, 1)$ three independent random variable. Then, for all $q \in \mathbb{N}$,*

$$\mathbb{E}\Big\{ \big[G + iU\big]^q \Big\} = \mathbb{E}\Big\{ \big[\sqrt{\sigma^2 - 1}\ \epsilon\big]^q \Big\}$$

Then, iU+G behaves like a Gaussian distribution with variance negative if $\sigma^2 < 1$. Now, if $\sigma^2 > 1$, we find a usual Gaussian distribution. If $\sigma^2 = 1$, the probability is concentrated in 0.

To use the negative variances in order to solve our problem, we will use the following property which is a consequence of proposition 4.1.2 : cf [9].

Proposition 4.1.3 *Let $U \sim N(0,1)$ be a random variable independent with $(\delta_1, \delta_2, ..., \delta_p)$ and $(\epsilon_1, \epsilon_2, ..., \epsilon_p)$. Then, for all $q \in \mathbb{N}$,*

$$\mathbb{E}\Big\{\Big(\sum_{t=1}^{p}(\delta_t z_t + \sigma_t \epsilon_t) + isU\Big)^q\Big\} = \mathbb{E}\big\{(T_s)^q\big\} \ ,$$

where $T_s = \sum_{t=1}^{p} \delta_t (z_t + \sqrt{\sigma_t^2 - s^2}\epsilon_t)$.

Proof This proposition is proved by using equalities based on simple calculations : cf [9]. ■

4.2 First application to mixtures

4.2.1 Method

Then, we can use the previous property in order to find the various parameters σ, z_r and q_r. For example, if $\sigma_r = \sigma$ for r=1,2,....,p, $\mathbb{E}\big\{[Y + i\sigma U]^q\big\} = \mathbb{E}\big\{S^q\big\}$. Therefore if we choose $s = \sigma$, $Y + isU$ has the same moments as $S = \sum_{r=1}^{p} \delta_r z_r$. Then its variance of order p will be equal to 0. Therefore, we can find $s = \sigma$ by studying the variances of order p of isU + Y. When σ is obtained, it suffices to calculate the parameters of the Gauss Jacobi quadrature which will be the associated the z_r's and q_r's.

Then, we can use the negative variances of Gaussian distributions in order to obtain the parameters of the mixture. Indeed, it is easy to prove the following proposition.

Proposition 4.2.1 *We assume that $\sigma_r^2 = \sigma^2$ for all $r \in \{1, 2, ..., p\}$. Then, $Y = \sum_{r=1}^{p} \delta_r z_r + \sigma\epsilon_0$ where $\epsilon_0 \sim N(0,1)$.*

Then, it is easy to find the parameters of this mixture by using the variances of higher order.

Indeed, assume that we want to estimate the z_t's and q_t's. We can estimate the moments of S by using negative variances because

$$\mathbb{E}\Big\{\Big(\sum_{t=1}^{p}(\delta_t z_t + \sigma\epsilon_t) + isU\Big)^q\Big\} = \mathbb{E}\Big\{\Big(\sum_{t=1}^{p}\delta_t\big[z_t + \sqrt{\sigma^2 - s^2}\ \epsilon_t\big]\Big)^q\Big\} \ .$$

It is thus enough to find s such that $\sigma^2 = s^2$: in this case, the moments of $\sum_{t=1}^{p}(\delta_t z_t + \sigma\epsilon_t) + isU$ have a variance of order p which is equal to zero. Indeed, it is the variance of order p of S which is concentrated in p distinct points.

Then, to find s such that $\sigma^2 = s^2$ is easy. Indeed, if we take s increasing from 0, $\sum_{t=1}^{p}(\delta_t z_t + \sigma\epsilon_t) + isU$ has the same moments as the real random variable $T_s = \sum_{t=1}^{p}(z_t + \sqrt{\sigma - s^2}\ \epsilon_t)$. We can then calculate the variance of order p of T_s. When it vanishes, we know that s is reached. Then we know the parameters searched $s = \sigma$ and also the z_r's and q_r's by using Gauss Jacobi quadrature.

Concretely, if we take s more and more large, variance of order p decreases to 0 and then becomes negative. Then, it is not difficult to find $s = \sigma$ by calculating the variance of order p for any value of s.

Example In figure 4.1, The blue curve represents the real part of the variance of order 3 of a sample of X + isU when X is a Gaussian mixture of components centered in -2,0,3, of standard deviation 1,1,1 of weight 1/6, 2/6, 3/6 when s varies from 0 to 1. The red curve represents the

imaginary part. We see that the variance vanishes near s = 1.

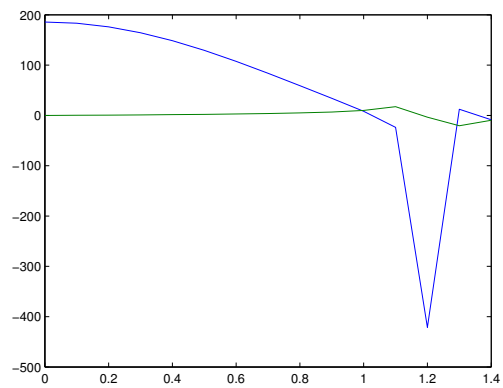

Figure 4.1: Variance of order 3 of T_s

To calculate σ, we'll take a curve defined over a smaller interval near the point where $s = \sigma$. In the figure 4.2, then we represent the same curves for intervals much smaller. This allows to conclude that $s_0 = \sigma = 0.996$.

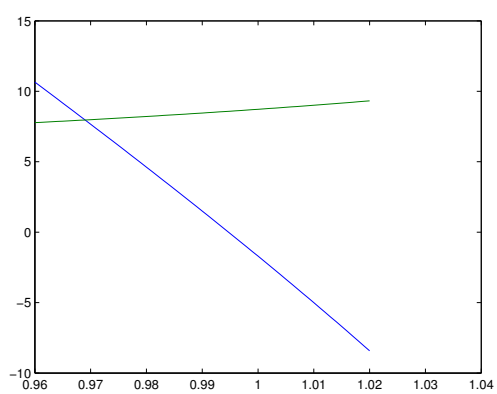

Figure 4.2: Variance of order 3 of T_s

With the moments, we can then estimate the parameters of the Gauss Jacobi quadrature.

We estimate the zeros by -1.9735, 0.0653, 3.0232 (instead of -2,0,3) with weights 0.1776, 0.3274, 0.4950 (instead of 0.1667, 0.3333, 0.5000).

Then, one thinks at an application : higher order regression (cf [9]). For example, if we have the development of real estate prices resulting from two different regions, we shall obtain the concentration around two curves. We shall treat this problem later.

4.3 Second application to mixtures

4.3.1 Presentation

Now we apply Proposition 4.1.3 when the σ_r 's are different : if $s^2 = Min_{r=1,\dots,p}(\sigma_r^2)$, $T_s = \sum_{t=1}^{p}(\delta_t z_t + \sigma_t \epsilon_t) + isU$ has the same moments as a Gaussian mixture with p components, including one degenerate, i.e. with a point of nonzero measure in z_{t_0} where $\sigma_{t_0}^2 = Min_{r=1,\dots,p}(\sigma_r^2)$: $P\{T_s = z_{t_0}\} = q_t > 0$. So we find $\sigma_{t_0}^2$ by applying a method of the same type as that described in chapter 3.

After we removed the points corresponding to the Gaussian component of the mixture and we start again.

4.3.2 Example

One chooze a sample with size 100000 of a Gaussian mixture with components centered in -3,0,2, with standard deviation 0,251,0.52,1, and with weights 0.4, 0.3, 0.3. (cf figure 4.3 : graph of the function hist of Matlab).

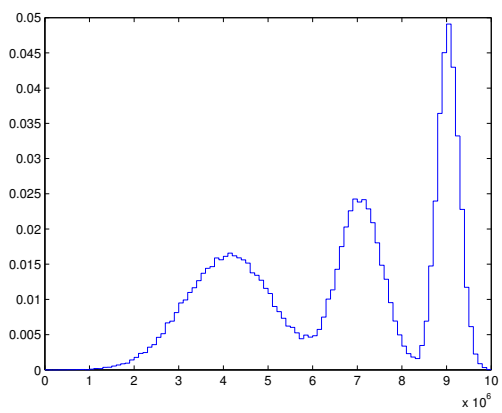

Figure 4.3: Function hist of Gaussian mixture with components centered in -3,0,2

4.3.3 Calculation of the first standard deviation

The first standard deviation which we can know by using the Gauss Jacobi quadrature is the smallest standard deviation.

Now, for a sample, this is only estimation obtained thanks to a complex random variable. Then, generally all the moments are not real and we can take the real part since the imaginary part of the moments tend to 0 as the sample size tends to ∞.

Therefore one estimates the moments and then the elements of the Gauss Jacobi quadrature with different s. In this report, we'll do it for s = 1/5 and s = 1/4.

For sake of simplicity, in this section we denote always by $\lambda_{j,s}$ and $x_{j,s}$ the weights and the zeros of the Gauss Jacobi quadrature associated with the sample of the random variable Y + isU, that is to say that the theoretical probabilities have the same moments as $\sum_{r=1}^{p} \delta_r \big(z_r + \sqrt{\sigma_r^2 - s^2}\ \epsilon_r\big)$.

Study of s=1/4 For s=1/4, for the zeros close to 2 and the associated weights we have the following table [1].

j	x_{j,j_0}	x_{j,j_0+1}	λ_{j,j_0}	λ_{j,j_0+1}
3	1.8229	0	0.3899	0
4	1.9695	0	0.3203	0
5	1.9932	0	0.3062	0
6	1.9991	0	0.3017	0
7	2.0010	0	0.2999	0
8	2.0014	0	0.2994	0
9	2.0014	0	0.2994	0
10	2.0658	2.0143	0.3533	0.0912
11	1.9053	2.0012	0.0000	0.2996
12	2.0019	0	0.2987	0
13	2.2172	1.9999	0.0000	0.3007
14	2.0062	0	0.2922	0
15	2.0691	1.9616	0.0942	0.2084
16	2.3039	1.9923	0.0048	0.2967

Study of s=1/5 Of course, in order to obtain the good value of s, we have to repeat this calculations several times. For example, if s=1/5, for the zeros close to 2 and the associated points, we have the following table.

j	x_{j,j_0}	x_{j,j_0+1}	λ_{j,j_0}	λ_{j,j_0+1}
3	1.8420	0	0.3860	0
4	2.0039	0	0.3104	0
5	2.0427	0	0.2874	0
6	2.0721	0	0.2632	0
7	2.1145	1.5975	0.2128	0.1126
8	2.1680	1.8255	0.1415	0.1678
9	2.2072	1.9006	0.0976	0.2040
10	2.2346	1.9395	0.0729	0.2217
11	2.0854	1.7481	0.1950	0.1016
12	2.1919	1.8889	0.1099	0.1913
13	2.2038	1.9050	0.0986	0.2006
14	2.2019	1.9019	0.1005	0.1994
15	2.2042	1.9056	0.0982	0.2009
16	2.2181	1.9230	0.0861	0.2100

By Stieljes Markov's inequality, we see that there is a weight of about 0.3 concentrated around 2. Indeed, let λ_{j,j_0} and λ_{j,j_0+1} and x_{j,j_0}, x_{j,j_0+1}, $x_{j,j_0} < 2 < x_{j,j_0+1}$, be the weights and the zeros associated around 2. Then, $P\{G_3 \in [2-a, 2+a]\} \leq \lambda_{j,j_0} + \lambda_{j,j_0+1}$ where G_3 represents the third Gaussian component when a is small. Now there is no perfectly correct point in order to identify it. We see only that it seems that there is a concentration close to 2.

This is normal because s is not well chosen. Indeed, at this point x=2, we have a Gaussian distribution with a variance of about $1/16 - 1/25 = 0.225$. This means that the probability is mainly concentrated on an interval of length 0.2 around 2. In order to be sure there are no points of nonzero measure, we should therefore consider the P_j where j is greater than 16. But the fact that there is no concentration points x_{js} seems enough accurate in order to indicate that for s = 1/5, there is no point of nonzero measure.

[1] for completetables, cf [9]

Conclusion Finally, we see a little difference between the results s = 1/5 and s = 1/4 if we confine ourselves to use the Markov Stieljes inequality. In order that the result appears more clearly, we must in fact take orthogonal polynomials of degree greater than 16.

On the other hand, the number of weights close to 0.3 concentrated in a single zero near 2 is much more important. In the simulations made otherwise, this is what is significant.

Other values of s By studying several values of s, we can determine about which value of s corresponds itself to a point of concentration. We can thus, by being located near this value, to find towards which points and which weights it is necessary to concentrate the researches as above in example 4.2.1 when $\sigma_r = \sigma$. We find then rather easily, the zero and the weight by proceeding as above by making a graph.

Estimation

By repeating many times this type of calculation we find that the best estimation for the concentration point is $z_3 = 2.02 \pm 0.01$. The calculation of s is a little more imprecise: we find $\sigma_3 = 0.25 \pm 0.02$. About the weight, it is estimated to be close to $q_3 = 0.3 \pm 0.03$.

4.3.4 Deletion of the first Gaussian component

We therefore estimated the parameters of the first Gaussian component. To find the second using the same technique, one must first eliminate points coming from the first component. There are various techniques possible. One choose to use a simple technique by eliminating $N_3 = nq_3$ points where n = 100000 is the size of sample.

Indded, we know that there is about a point of the sample coming from that component in a partition of $\mathbb{R}$ in N_3 intervals whose the Gaussian probability of each is $1/N_3$. One can then choose to remove the point nearest point of the middle of each interval. We did it in [9].

The result obtained is provided in Figure 4.4. In this figure 4.4 : for a sample of n = 100000, we have delete $nq_3 = 29500$ points supposed to belong to the first Gaussian component. The graph is the function histogram of Matlab (hist).

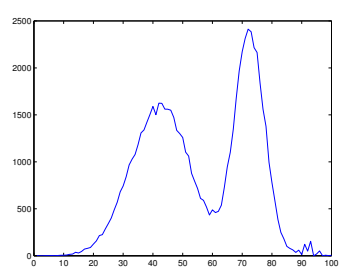

Figure 4.4: Function hist of the deletion of the normal component centered in 2

Note that these graphs can be used to better determine the weight, that is to say the size of the sample which has to be removed. Indeed, the curve has peaks on his right incompatible with a Gaussian mixture with two components.

Then, one can try another weight close to the estimation which we got. For example, if one removes only 30150 point we have the graph of Figure 4.5.

This kind of results can eventually allow us to refine the choice of s.

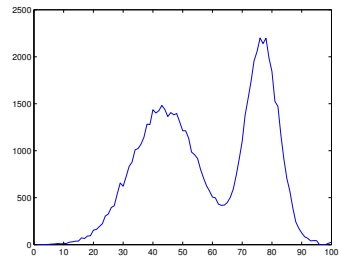

Figure 4.5: Function hist of the deletion of 30150 elements of the normal component centered in 2

4.3.5 Estimation of the Other Gaussian components

Then, one proceeds by the same way and the parameters of other gaussian components are obtained. Then, we found

$z_3 = 2.02 \pm 0.01$, $z_2 = 0.03 \pm 0.01$, $z_1 = -3.02 \pm 0.01$.
$\sigma_3 = 0.25 \pm 0.01$, $\sigma_2 = 0.50 \pm 0.01$, $\sigma_1 = 1.01 \pm 0.02$.
$q_3 = 0.3 \pm 0.03$, $q_2 = 0.3 \pm 0.03$, $q_1 = 0.4 \pm 0.02$.

4.3.6 Conclusion

So we have relatively good results especially when you consider that these results can be refined by a more detailed study of properties of the Gauss Jacobi quadrature.

Of course, one may want to compare this method with other methods used in order to find the Gaussian components, that is to say, essentially, the EM algorithm (cf [11]) : by applying it to the sample which we have just estimate, it was found.

$z_3 = 2.1$, $z_2 = -0.2$ et $z_1 = -2.99$.
$\sigma_3 = 0.26$, $\sigma_2 = 0.49$, $\sigma_1 = 0.98$.
$q_3 = 0.28$, $q_2 = 0.31$, $q_1 = 0.39$.

So, in this case, our method gives result best than those of the EM algorithm.

Appendix A

Standardized Variances

We are going to study examples of standardized variances $\hat{\sigma}_j^2$: at first normal centered laws, then Gaussian mixtures, then Gaussian mixtures truncated.

Finally, we shall see that if we want that the value of the variances can measure the concentration as soon as it is obtained, it is necessary to proceed to the second standardization by centering the random variables X or by imposing by translation that they are positive.

A.1 Centered Gaussian distributionss

We recall that, by proposition 1.2.2, the variance of a standard Gaussian distribution is

$$\left(\sigma_j^{01}\right)^2 = j! \ .$$

Therefore, for j=1,2,3,4, σ_j^2 = 1,2,6,24.

On the other hand $||x^j||^2 = \int x^{2j} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \frac{2j!}{2^j j!}$.

Therefore, for j=1,2,3,4, $||x^j||^2 = \frac{2!}{2^1 1!} = 1,\ \frac{4!}{2^2 2!} = 3,\ \frac{6!}{2^3 3!} = 15,\ \frac{8!}{2^4 4!} = 105$.
i.e. , $\frac{\sigma_j^2}{||x^j||^2} = \frac{1}{1},\ \frac{2}{3},\ \frac{6}{15},\ \frac{24}{105}$.
i.e., $\frac{\sigma_j^2}{||x^j||^2} = 1,\ 0.66,\ 0.4,\ 0.2286$.

A.2 Gaussian mixtures

We study Gaussian mixtures. We use notations 4.1.1 : $Y = \sum_{r=1}^{p} \delta_r Z_r$, where $Z_r - z_r = \sigma_r^G \epsilon_r \sim N(0, (\sigma_r^G)^2)$, and $P(\delta_r = 1) = q_r$.

We denote these mixtures by "MixtGauss ([Expectation],[Weights],[Variances])" where $[Expectation] = [z_1, ..., z_p]$, $[Weights] = [q_1, ..., q_p]$, and $[Variances] = [(\sigma_1^G)^2, ...(\sigma_p^G)^2]$.

One denotes again standardized variances of order 2,3,4 associated to these mixtures by $\hat{\sigma}_j^2 = \frac{\sigma_j^2}{||x^{2j}||}$: $\hat{\sigma}_2^2$, $\hat{\sigma}_3^2$, $\hat{\sigma}_4^2$.

Example 1 : Variance of order 2

In this example we study the mixture MixtGauss([-0.5,0.5],[1/2,1/2],[0.5,1]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.6276$, $\hat{\sigma}_3^2 = 0.3295$, $\hat{\sigma}_4^2 = 0.1596$.

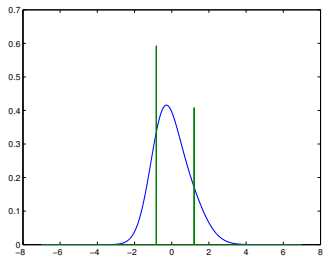

Figure A.1: $\hat{\sigma}_2^2 = 0.6276$.

Example 2 : Variance of order 2

In this example we study the mixtures MixtGauss([-1,1],[1/2,1/2],[0.5,1]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.5412$, $\hat{\sigma}_3^2 = 0.2833$, $\hat{\sigma}_4^2 = 0.1291$.

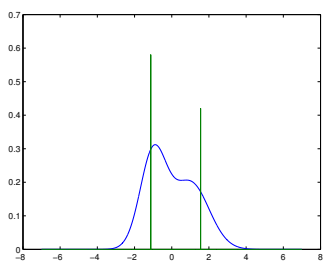

Figure A.2: $\hat{\sigma}_2^2 = 0.5412$

Example 3 : Variance of order 2

In this example we study the mixtures MixtGauss(([-1,2],[1/2,1/2],[0.5,1]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.3200$, $\hat{\sigma}_3^2 = 0.1350$, $\hat{\sigma}_4^2 = 0.0459$.

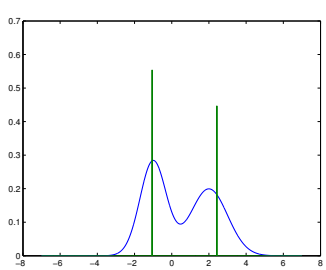

Figure A.3: $\hat{\sigma}_2^2 = 0.3200$

Example 4 : Variance of order 2

In this example we study the mixtures MixtGauss([-1,2],[1/4,3/4],[0.5,1]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.2312$, $\hat{\sigma}_3^2 = 0.0816$, $\hat{\sigma}_4^2 = 0.0302$.

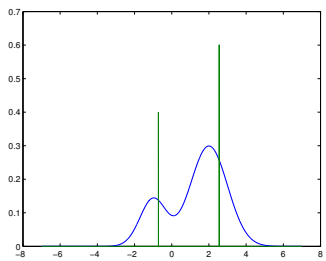

Figure A.4: $\hat{\sigma}_2^2 = 0.2312$

Example 5 : Variance of order 2

In this example we study the mixtures MixtGauss([-1,2],[3/4,1/4],[0.5,1]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.4572$, $\hat{\sigma}_3^2 = 0.2038$, $\hat{\sigma}_4^2 = 0.0724$.

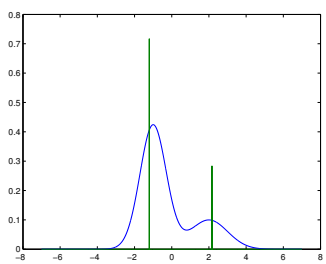

Figure A.5: $\hat{\sigma}_2^2 = 0.4572$

Example 6 : Variance of order 2

In this example we study the mixtures MixtGauss([-1,2],[1/4,3/4],[0.5,0.3]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.1810$, $\hat{\sigma}_3^2 = 0.0917$, $\hat{\sigma}_4^2 = 0.0351$.

Example 7 : Variance of order 2

In this example we study the mixtures MixtGauss(([-1,2],[1,1],[0.5,0.3]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.2745$, $\hat{\sigma}_3^2 = 0.1529$, $\hat{\sigma}_4^2 = 0.00596$.

Example 8 : Variance of order 2

In this example we study the mixtures MixtGauss([-2,2],[1,1],[0.5,0.3]) and the variance of order 2.

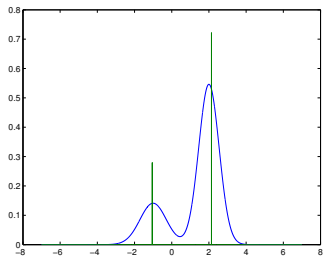

Figure A.6: $\hat{\sigma}_2^2 = 0.1810$

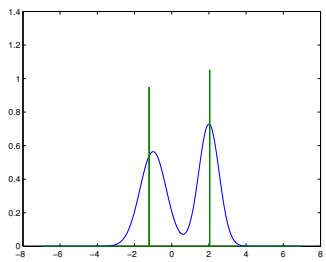

Figure A.7: $\hat{\sigma}_2^2 = 0.2745$

We obtain $\hat{\sigma}_2^2 = 0.2554$, $\hat{\sigma}_3^2 = 0.1763$, $\hat{\sigma}_4^2 = 0.0621$.

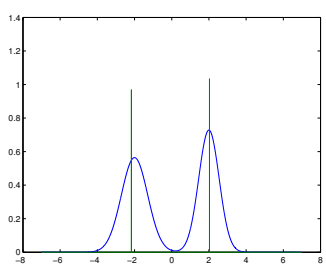

Figure A.8: $\hat{\sigma}_2^2 = 0.2554$

Example 9 : Variance of order 2

In this example we study the mixtures MixtGauss([-3,2],[1,1],[0.5,0.3]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.1553$, $\hat{\sigma}_3^2 = 0.0925$, $\hat{\sigma}_4^2 = 0.0210$.

Example 10 : Variance of order 2

In this example we study the mixtures MixtGauss([-3,2],[1,1],[0.1,0.3]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.0914$, $\hat{\sigma}_3^2 = 0.0544$, $\hat{\sigma}_4^2 = 0.0125$.

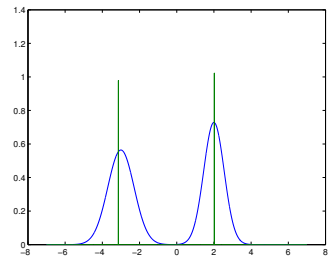

Figure A.9: $\hat{\sigma}_2^2 = 0.1553$

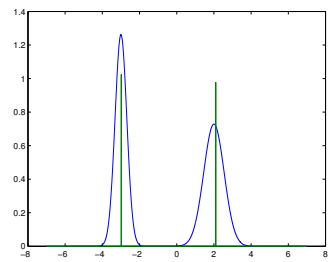

Figure A.10: $\hat{\sigma}_2^2 = 0.0914$

Example 11 : Variance of order 3

In this example we study the mixtures MixtGauss([-3,0,2],[1,1,1],[0.1,0.3,0.3]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.3077$, $\hat{\sigma}_3^2 = 0.0700$, $\hat{\sigma}_4^2 = 0.0377$.

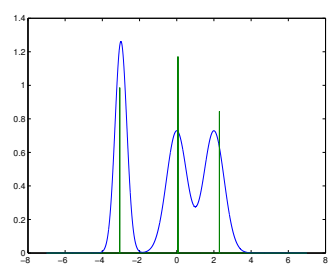

Figure A.11: $\hat{\sigma}_3^2 = 0.0700$

Example 12 : Variance of order 3

In this example we study the mixtures MixtGauss(([-3,0,2],[1/3,1/3,1/3],[0.1,0.3,0.05]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.2520$, $\hat{\sigma}_3^2 = 0.0382$, $\hat{\sigma}_4^2 = 0.0148$.

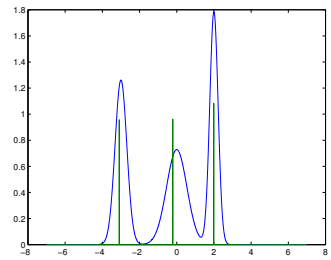

Figure A.12: $\hat{\sigma}_3^2 = 0.0382$

Example 13 : Variance of order 3

In this example we study the mixtures MixtGauss([-3,0,2],[0.25,0.32,1-0.32-0.25],[0.5,0.3,0.2]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.3711$, $\hat{\sigma}_3^2 = 0.1104$, $\hat{\sigma}_4^2 = 0.0440$.

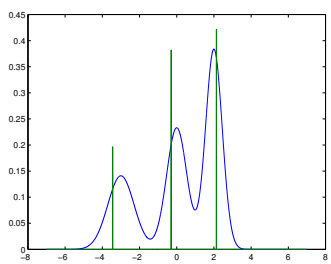

Figure A.13: $\hat{\sigma}_3^2 = 0.1104$

Example 14 : Variance of order 3

In this example we study the mixtures MixtGauss(([-3,0,2],[0.25,0.32,1-0.32-0.25],[0.5,0.7,0.9]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.4736$, $\hat{\sigma}_3^2 = 0.1988$, $\hat{\sigma}_4^2 = 0.1049$.

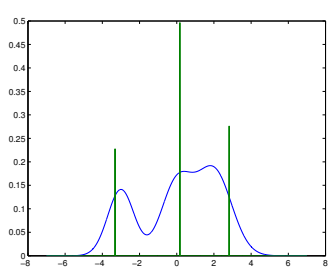

Figure A.14: $\hat{\sigma}_3^2 = 0.1988$

Example 15 : Variance of order 3

In this example we study the mixtures MixtGauss([-3,0,2],[0.25,0.32,1-0.32-0.25],[0.1,0.3,0.3]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.2939$, $\hat{\sigma}_3^2 = 0.0434$, $\hat{\sigma}_4^2 = 0.0190$.

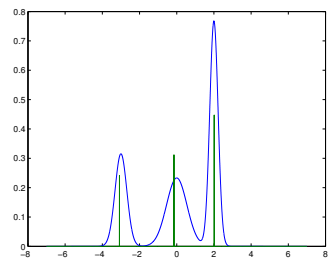

Figure A.15: $\hat{\sigma}_3^2 = 0.0434$

Example 16 : Variance of order 4

In this example we study the mixtures MixtGauss([-3,0,2],[0.25,0.32,1-0.32-0.25],[1.1,1.2,0.9]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.4801$, $\hat{\sigma}_3^2 = 0.2195$, $\hat{\sigma}_4^2 = 0.0971$.

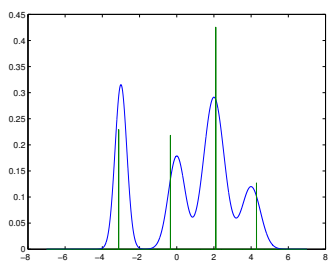

Figure A.16: $\hat{\sigma}_4^2 = 0.0971$

Example 17 : Variance of order 4

In this example we study the mixtures MixtGauss([-3,0,2,4],[0.25,0.2,0.4,1-0.2-0.4-0.25],[0.1,0.2,0.3,0.25]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.4089$, $\hat{\sigma}_3^2 = 0.1018$, $\hat{\sigma}_4^2 = 0.0296$.

Example 18 : Variance of order 4

In this example we study the mixtures MixtGauss([-3,0,2,4],[0.25,0.2,0.4,1-0.2-0.4-0.25],[0.1,0.2,0.3,0.25]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.3084$, $\hat{\sigma}_3^2 = 0.0666$, $\hat{\sigma}_4^2 = 0.0091$.

Example 19 : Variance of order 4

In this example we study the mixtures MixtGauss([-3,0,2,4],[0.25,0.2,0.4,1-0.2-0.4-0.25],[0.1,0.05,0.15,0.07]) and the variance of order 4.

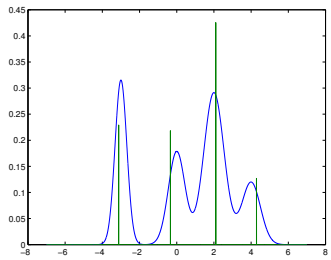

Figure A.17: $\hat{\sigma}_4^2 = 0.0296$

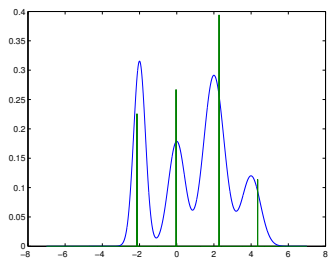

Figure A.18: $\hat{\sigma}_4^2 = 0.0091$

We obtain $\hat{\sigma}_2^2 = 0.3057$, $\hat{\sigma}_3^2 = 0.0622$, $\hat{\sigma}_4^2 = 0.0091$.

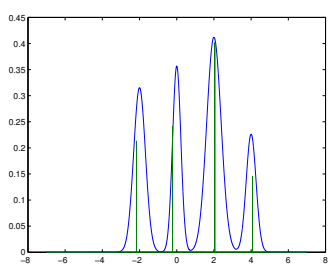

Figure A.19: $\hat{\sigma}_4^2 = 0.0091$

Example 20 : Variance of order 4

In this example we study the mixtures MixtGauss([-3,0,2,4],[0.25,0.2,0.4,1-0.2-0.4-0.25],[0.1,0.05,0.15,0.07]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.1986$, $\hat{\sigma}_3^2 = 0.0396$, $\hat{\sigma}_4^2 = 0.0095$.

A.3 Truncated gaussian mixtures

One study again Gaussian mixtures $Y = \sum_{r=1}^{p} \delta_r Z_r$. But we truncate them : we restrict them on [0,1]. We obtain a density f_0 which is equal to 0 anywhere other than on [0,1]. Then we

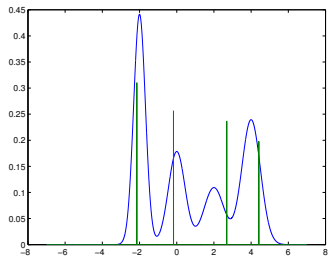

Figure A.20: $\hat{\sigma}_4^2 = 0.0095$

multiply this density truncated by a coefficient c_0 so as to have always a density of probability : i.e. $\int_{x=0}^{1} c_0 f_0(x)dx = 1$.

We denote these truncated mixtures by "MixtGauss Trunc([Expectation],[Weights],[Variances])" where $[Expectation] = [z_1, ..., z_p]$, , $[Weights] = [q_1, ..., q_p]$, and $[Variances] = [(\sigma_1^G)^2, ...(\sigma_p^G)^2]$.

We denote again standardized variances by $\hat{\sigma}_j^2 = \frac{\sigma_j^2}{||x^{2j}||}$.

A.3.1 First type of graphs

On the graphs of examples 21-26, the weigths of Gauss Jacobi are multiplied by 10 so as to have more readable graphs.

Example 21 : Variance of order 4

In this example we study the mixture MixtGaussTrunc([0.1,0.3,0.55,0.8],[0.35,0.2,0.15,1-0.2-0.15-0.35],[0.0003,0.0002,0.0003,0.00025]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.0190$, $\hat{\sigma}_3^2 = 0.0008232$, $\hat{\sigma}_4^2 = 0.00002$.

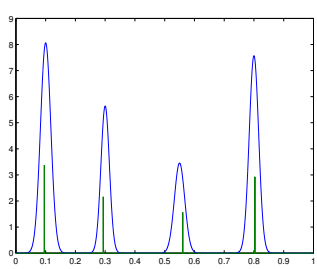

Figure A.21: $\hat{\sigma}_4^2 = 0.00002$

Example 22 : Variance of order 4

In this example we study the mixture MixtGaussTrunc([0.1,0.3,0.55,0.8],[0.45,0.2,0.25,1-0.2-0.25-0.45],[0.0003,0.0002,0.0003,0.00025]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.0352$, $\hat{\sigma}_3^2 = 0.0023$, $\hat{\sigma}_4^2 = 0.000042$.

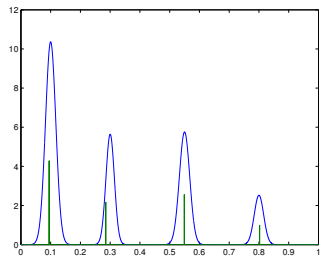

Figure A.22: $\hat{\sigma}_4^2 = 0.000042$

Example 23 : Variance of order 4

In this example we study the mixture MixtGaussTrunc([0.1,0.3,0.55,0.8],[0.45,0.2,0.25,1-0.2-0.25-0.45],[0.0001,0.0004,0.0003,0.00025]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.0346$, $\hat{\sigma}_3^2 = 0.0021$, $\hat{\sigma}_4^2 = 0.000025$.

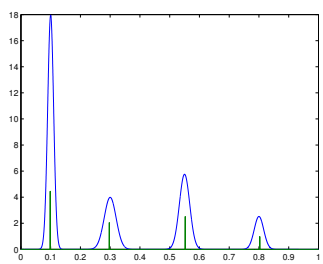

Figure A.23: $\hat{\sigma}_4^2 = 0.000025$

Example 24 : Variance of order 4

In this example we study the mixture MixtGaussTrunc(0.1,0.3,0.55,0.8],[0.45,0.2,0.25,1-0.2-0.25-0.45],[0.001,0.004,0.003,0.0025]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.0380$, $\hat{\sigma}_3^2 = 0.0032$, $\hat{\sigma}_4^2 = 0.000237$.

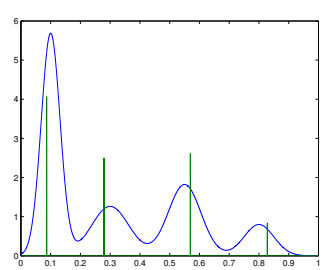

Figure A.24: $\hat{\sigma}_4^2 = 0.000237$

Example 25 : Variance of order 4

In this example we study the mixture MixtGaussTrunc([0.1,0.3,0.55,0.8],[0.25,0.2,0.35,1-0.2-0.35-0.25],[0.001,0.004,0.003,0.0025]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.0274$, $\hat{\sigma}_3^2 = 0.0019$, $\hat{\sigma}_4^2 = 0.000148$.

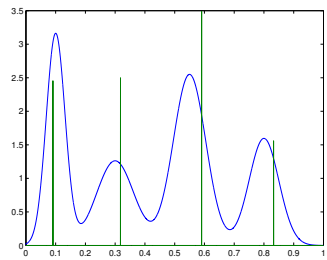

Figure A.25: $\hat{\sigma}_4^2 = 0.000148$

Example 26: Variance of order 4

In this example we study the mixture MixtGaussTrunc([0.1,0.3,0.55,0.8],[0.25,0.2,0.35,1-0.2-0.35-0.25],[0.001,0.004,0.003,0.0025]) and the variance of order 4.

We obtain $\hat{\sigma}_2^2 = 0.0316$, $\hat{\sigma}_3^2 = 0.0032$, $\hat{\sigma}_4^2 = 0.000302$.

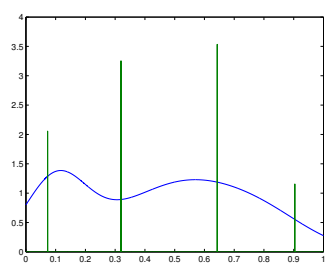

Figure A.26: $\hat{\sigma}_4^2 = 0.000302$

A.3.2 Second type of graphs

On the graphs of examples 27-39, the weigths of Gauss Jacobi are multiplied by 5 so as to have more readable graphs.

Example 27 : Variance of order 3

In this example we study the mixture MixtGaussTrunc([0.1,0.3,0.8],[0.2,0.35,1-0.2-0.35],[0.002,0.005,0.01]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.0183$, $\hat{\sigma}_3^2 = 0.0021$, $\hat{\sigma}_4^2 = 0.0001243$.

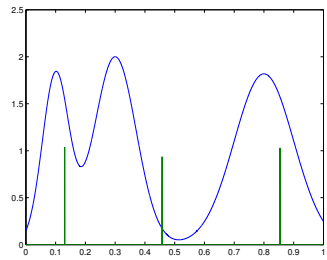

Figure A.27: $\hat{\sigma}_3^2 = 0.0021$

Example 28 : Variance of order 3

In this example we study the mixture MixtGaussTrunc([0.1,0.45,0.8],[0.2,0.35,1-0.2-0.35],[0.002,0.005,0.01]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.0232$, $\hat{\sigma}_3^2 = 0.0016$, $\hat{\sigma}_4^2 = 0.0001399$.

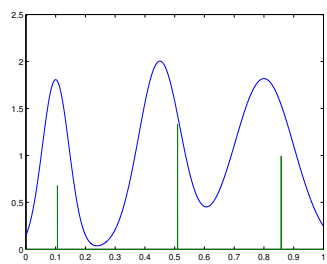

Figure A.28: $\hat{\sigma}_3^2 = 0.0016$

Example 29 : Variance of order 2

In this example we study the mixture MixtGaussTrunc([0.1,0.65,0.8],[0.2,0.35,1-0.2-0.35],[0.002,0.005,0.01]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.0144$, $\hat{\sigma}_3^2 = 0.00075$, $\hat{\sigma}_4^2 = 0.000116$.

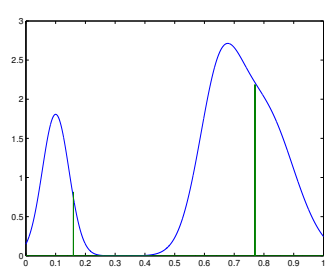

Figure A.29: $\hat{\sigma}_2^2 = 0.0144$

Example 30 : Variance of order 3

In this example we study the mixture MixtGaussTrunc([0.1,0.65,0.8],[0.2,0.35,1-0.2-0.35],[0.002,0.001,0.002]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.0089$, $\hat{\sigma}_3^2 = 0.0004456$, $\hat{\sigma}_4^2 = 0.0000521$.

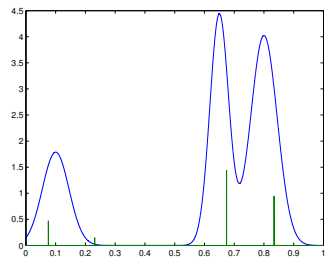

Figure A.30: $\hat{\sigma}_3^2 = 0.0004456$

Example 31 : Variance of order 3

In this example we study the mixture MixtGaussTrunc([0.1,0.4,0.8],[0.4,0.35,1-0.4-0.35],[0.002,0.001,0.002]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.0341$, $\hat{\sigma}_3^2 = 0.0011$, $\hat{\sigma}_4^2 = 0.00018$.

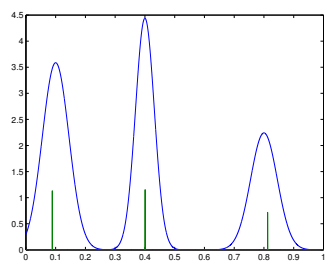

Figure A.31: $\hat{\sigma}_3^2 = 0.0011$

Example 32 : Variance of order 3

In this example we study the mixture MixtGaussTrunc([0.1,0.4,0.8],[0.4,0.35,1-0.4-0.35],[0.006,0.007,0.005]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.0345$, $\hat{\sigma}_3^2 = 0.0024$, $\hat{\sigma}_4^2 = 0.000283$.

Example 33 : Variance of order 3

In this example we study the mixture MixtGaussTrunc([0.1,0.4,0.8],[0.4,0.35,1-0.4-0.35],[0.01,0.013,0.011]) and the variance of order 3.

We obtain $\hat{\sigma}_2^2 = 0.0358$, $\hat{\sigma}_3^2 = 0.0031$, $\hat{\sigma}_4^2 = 0.000312$.

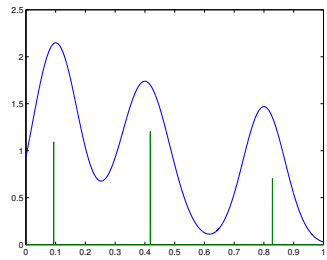

Figure A.32: $\hat{\sigma}_3^2 = 0.0024$

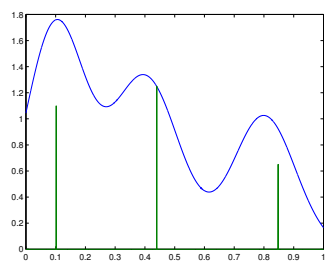

Figure A.33: $\hat{\sigma}_3^2 = 0.0031$

Example 34 : Variance of order 2

In this example we study the mixture MixtGaussTrunc([0.1,0.3,0.8],[0.4,0.35,1-0.4-0.35],[0.01,0.013,0.011]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.0337$, $\hat{\sigma}_3^2 = 0.0032$, $\hat{\sigma}_4^2 = 0.000288$.

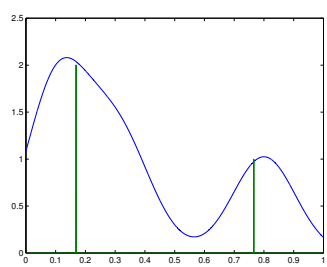

Figure A.34: $\hat{\sigma}_2^2 = 0.0337$

Example 35 : Variance of order 2

In this example we study the mixture MixtGaussTrunc([0.2 0.8],[0.4,0.6],[0.01,0.013]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.0126$, $\hat{\sigma}_3^2 = 0.0016$, $\hat{\sigma}_4^2 = 0.000102$.

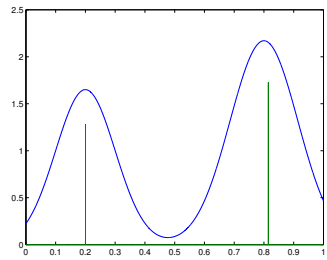

Figure A.35: $\hat{\sigma}_2^2 = 0.0126$

Example 36 : Variance of order 2

In this example we study the mixture MixtGaussTrunc([0.3 0.8],[0.4,0.6],[0.01,0.013]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.0126$, $\hat{\sigma}_3^2 = 0.0010$, $\hat{\sigma}_4^2 = 0.000072$.

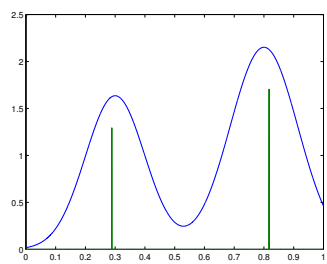

Figure A.36: $\hat{\sigma}_2^2 = 0.0126$

Example 37 : Variance of order 2

In this example we study the mixture MixtGaussTrunc([0.3 0.8],[0.3,0.7],[0.03,0.013]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.0107$, $\hat{\sigma}_3^2 = 0.0011$, $\hat{\sigma}_4^2 = 0.000089$.

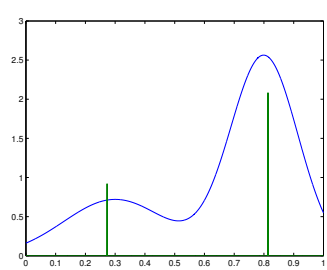

Figure A.37: $\hat{\sigma}_2^2 = 0.0107$

Example 38 : Variance of order 2

In this example we study the mixture MixtGaussTrunc([0.3 0.8],[0.3,0.7],[0.03,0.04]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.0150$, $\hat{\sigma}_3^2 = 0.0014$, $\hat{\sigma}_4^2 = 0.000118$.

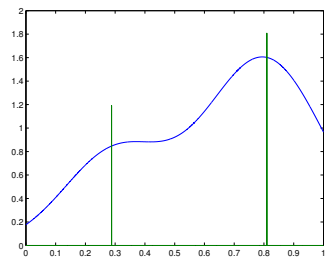

Figure A.38: $\hat{\sigma}_2^2 = 0.0150$

Example 39 : Variance of order 2

In this example we study the mixture MixtGaussTrunc([0.3 0.8],[0.45,0.55],[0.03,0.02]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.0164$, $\hat{\sigma}_3^2 = 0.0017$, $\hat{\sigma}_4^2 = 0.0001178$.

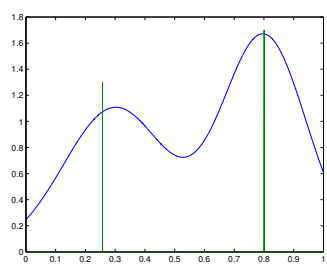

Figure A.39: $\hat{\sigma}_2^2 = 0.0164$

A.4 Comparison of the various results

We could wonder because the variances of the mixture truncated in [0,1] are smaller than those of the variances of Gaussian classic mixtures studied in section A.2. Indeed we could think that, according to proposition 1.3.8, these standardized variances should not change because $\sigma_j(\alpha X)^2 = \alpha^{2j}\sigma_j^2$ and because $||(\alpha X)^{2j}|| = \alpha^{2j}||X^{2j}||$: therefore $\sigma_j(\alpha X)^2 = \sigma_j(X)^2$. But there is no error: it is indeed the case. We can see it by deleting the truncation: in that case, the standardized variances are much smaller than those of the section A.2.

In fact the difference comes from what it is the translation which changes because we put ourselves in [0,1] thus positive. In the case of translation, the variances (not standardized) do not change according to proposition 1.3.3. But, when we make a translation of a, in order to have the standardized variances, it is necessary to divide, not by $||X^{2j}||$, but by $||(X+a)^{2j}||$. And it is the moments $||(X+a)^{2j}||$ which change if a varies.

Finally, in order to have a complete standardization, it is necessary or to center the random variable X or to make it positive by a translation.

We can see this by developing the example 39.

Example 40 : Variance of order 2

We use again example 39, but we do not truncate the probability, i.e. we study the mixture MixtGauss([0.3 0.8],[0.45,0.55],[0.03,0.02]) and the variance of order 2.

We obtain $\hat{\sigma}_2^2 = 0.0261$, $\hat{\sigma}_3^2 = 0.0047$, $\hat{\sigma}_4^2 = 0.0008356$.

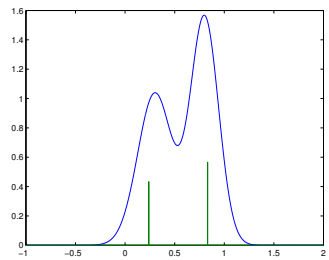

Figure A.40: $\hat{\sigma}_2^2 = 0.0261$

Example 41 : Variance of order 2

In this example, we use again example 40 where the variable X_{40} has the distribution of the mixture MixtGaussTrunc([0.3 0.8],[0.45,0.55],[0.03,0.02]), but we are interested by the variable $5X_{40}$ de loi MixtGauss(5[0.3 0.8],[0.45,0.55],25[0.03,0.02]).

We obtain $\hat{\sigma}_2^2 = 0.0261$, $\hat{\sigma}_3^2 = 0.0047$, $\hat{\sigma}_4^2 = 0.000829$.

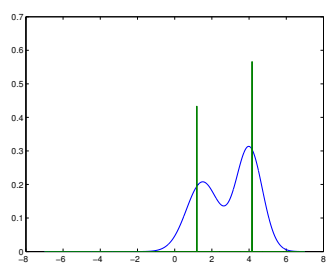

Figure A.41: $\hat{\sigma}_2^2 = 0.0261$

Example 42 : Variance of order 2

In this example we use again example 41 ($5X_{40}$) which we translate with a=2.5 : $5X_{40} - 2.5$, i.e. we study the mixture MixtGauss(5*[0.3 0.8]-2.5,[0.45,0.55],25*[0.03,0.02]).

We obtain $\hat{\sigma}_2^2 = 0.4515$, $\hat{\sigma}_3^2 = 0.2655$, $\hat{\sigma}_4^2 = 0.1339$.

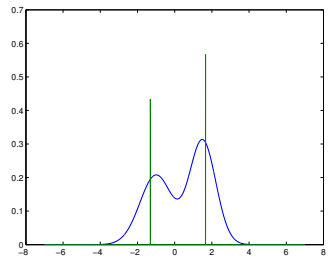

Figure A.42: $\hat{\sigma}_2^2 = 0.4515$

A.5 Conclusion

We thus see that the standardization $\frac{\sigma_j^2}{||x^{2j}||}$ gives a first way in order to be able to compare in a simple way the meaning of the value of the variances. But in fact, it is insufficient because if we make a translation, the standardized variances change, although they describe nevertheless the same type of concentration. Thus if we want to standardize better the variances, it seems that it is necessary to use or centered standardized variances $\hat{\sigma}_j^2$ either to impose on them to be positive with a very strong probability.

It seems that it is the good method. In that case, only with the value of the standardized variance of order j, we shall know generally if there is concentration or not near j distinct points.

Appendix B

Example of the decomposition of a Gaussian mixture

We study again Gaussian mixtures and we keep notations 4.1.1 : $Y = \sum_{r=1}^{p} \delta_r Z_r$, where $Z_r - z_r = \sigma_r^G \epsilon_r \sim N(0, (\sigma_r^G)^2)$, and $P(\delta_r = 1) = q_r$. In this chapter, we study a Gaussian mixture with p=2, $z_t = -2, 2$, $q_s = 0.3, 0.7$ and $(\sigma_s^G)^2 = 0.4, 0.2$.

The purpose thus is to find the various parameters of this Gaussian mixture by studying zeros and weights of the orthogonal polynomials associated in moments of Y+isU for different s .

According to method used in section 4.3, the first component which we have to detect is thus centered in $z_1 = 2$ with weights $q_1 = 0.7$ and variances $(\sigma_1^G)^2 = 2$.

To show how we can proceed in that case we shall give zeros and weights obtained for the moments of iUs+Y for $s = 0.4472$ ($s^2 = 0.2$), $s = 0.42$ and $s = 0.465$.

B.1 Study of the case $s = 0.4472$

In this section, we give zeros and obtained weights when we choose $s^2 = 0.2 = 0.4472^2$, i.e. the case where the moments of Y+isU are the same that those of a degenerate Gaussian mixture : because p=2, it is therefore the mixture of a single Gaussian variable centered in $z_2 = -2$ and of a point of measure not equal to 0 in $z_1 = 2$.

In this chapter, we keep the same notations as in chapters 3 and 4.

	1	2	3	4	5
x_{1s}	0.7999 + 0.0021i	0	0	0	0
x_{2s}	-2.0947 + 0.0035i	1.9785 + 0.0019i	0	0	0
x_{3s}	1.9984 + 0.0006i	-2.5543 + 0.0010i	-1.6664 - 0.0217i	0	0
x_{4s}	-6.4392 + 3.0067i	1.9987 - 0.0000i	-2.5021 - 0.0271i	-1.6170 - 0.0739i	0
x_{5s}	4.6792 + 3.3469i	2.0001 - 0.0011i	-2.8393 + 0.1739i	-2.0098 + 0.1257i	-0.8555 - 0.0184i
x_{6s}	3.7629 + 1.1921i	1.1409 - 3.2385i	1.9988 + 0.0007i	-2.7987 + 0.1451i	-2.0787 + 0.5164i
x_{7s}	-2.7246 - 8.6751i	4.3698 + 1.0096i	1.9993 - 0.0015i	-2.4209 + 1.9289i	-2.9775 + 0.0538i
x_{8s}	5.7651 - 8.0617i	-6.9879 - 4.4457i	4.3512 + 1.1518i	-1.9338 + 2.6954i	1.9996 - 0.0015i
x_{9s}	3.8350 + 8.1704i	3.8500 - 5.8970i	4.3871 + 0.8763i	-4.4901 - 3.2670i	1.9995 - 0.0017i
x_{10s}	-10.6214 + 3.3505i	3.1691 + 7.4505i	4.3514 - 4.5626i	-3.2789 - 4.3978i	4.2331 + 0.9140i

	6	7	8	9	10
x_{6s}	-1.6538 + 0.1866i	0	0	0	0
x_{7s}	-1.9770 + 0.0228i	-0.5424 - 0.7657i	0	0	0
x_{8s}	-2.9646 + 0.0811i	-2.0109 + 0.0324i	-0.6832 - 0.4678i	0	0
x_{9s}	1.8169 + 1.9957i	-2.9771 + 0.0321i	-2.0009 + 0.0290i	-0.4954 - 0.5376i	0
x_{10s}	1.9994 - 0.0016i	-2.2264 + 2.0314i	-2.9982 + 0.0582i	-2.0010 + 0.0263i	-0.5923 - 0.6366i

	1	2	3	4	5
λ_{1s}	1	0	0	0	0
λ_{2s}	0.2894 + 0.0006i	0.7106 - 0.0006i	0	0	0
λ_{3s}	0.7123 - 0.0001i	0.1624 - 0.0026i	0.1253 + 0.0027i	0	0
λ_{4s}	-0.0000 - 0.0036i	0.7183 + 0.0021i	0.2154 + 0.0159i	0.0663 - 0.0144i	0
λ_{5s}	-0.0026 - 0.0028i	0.7107 + 0.0226i	0.0471 - 0.0090i	0.2297 + 0.0813i	0.0151 - 0.0921i
λ_{6s}	0.0032 - 0.0116i	-0.0009 - 0.0052i	0.7016 + 0.0177i	0.0887 + 0.0479i	0.0420 + 0.0104i
λ_{7s}	-0.0000 + 0.0001i	0.0038 - 0.0064i	0.7078 + 0.0197i	0.0017 - 0.0218i	0.0753 - 0.0408i
λ_{8s}	-0.0000 - 0.0000i	0.0000 - 0.0001i	0.0030 - 0.0063i	0.0070 - 0.0075i	0.7073 + 0.0199i
λ_{9s}	-0.0000 + 0.0000i	-0.0001 - 0.0001i	0.0041 - 0.0058i	-0.0001 - 0.0005i	0.7079 + 0.0196i
λ_{10s}	-0.0000 - 0.0000i	-0.0000 + 0.0000i	-0.0002 - 0.0001i	-0.0005 - 0.0000i	0.0040 - 0.0066i

	6	7	8	9	10
λ_{6s}	0.1653 - 0.0591i	0	0	0	0
λ_{7s}	0.2000 + 0.1151i	0.0114 - 0.0659i	0	0	0
λ_{8s}	0.0558 - 0.0342i	0.2144 + 0.1030i	0.0125 - 0.0747i	0	0
λ_{9s}	0.0163 - 0.0149i	0.0622 - 0.0329i	0.2030 + 0.1087i	0.0067 - 0.0741i	0
λ_{10s}	0.7080 + 0.0192i	0.0067 - 0.0190i	0.0640 - 0.0367i	0.2095 + 0.1140i	0.0086 - 0.0710i

(j,k)	$x_{j,k}$	$\lambda_{j,k}$
(2,2)	-1.9785 + 0.0019i	0.7106 - 0.0006i
(3,1)	1.9984 + 0.0006i	0.7123 - 0.0001i
(4,2)	1.9987 - 0.0000i	0.7183 + 0.0021i
(5,2)	2.0001 - 0.0011i	0.7107 + 0.0226i
(6,3)	1.9988 + 0.0007i	0.7016 + 0.0177i
(7,3)	1.9993 - 0.0015i	0.7078 + 0.0197i
(8,5)	1.9996 - 0.0015i	0.7073 + 0.0199i
(9,5)	1.9995 - 0.0017i	0.7079 + 0.0196i
(10,6)	1.9994 - 0.0016i	0.7080 + 0.0192i

B.2 Study of the case $s = 0.42$

Now we give the zeros and the weights obtained when $s = 0.42$. By proposition 4.1.3, it is thus weights of a Gaussian mixture.

	1	2	3	4	5
x_{1s}	0.8012 + 0.0003i	0	0	0	0
x_{2s}	-2.0910 - 0.0005i	1.9896 - 0.0004i	0	0	0
x_{3s}	2.0250 - 0.0008i	-2.3581 - 0.0067i	-1.0526 - 0.0235i	0	0
x_{4s}	-2.7331 - 0.0329i	-1.7908 - 0.0137i	2.0724 + 0.0102i	1.5103 + 0.0818i	0
x_{5s}	2.0326 + 0.0769i	2.0005 + 0.3007i	-3.0517 - 0.0034i	-2.1117 + 0.0709i	-1.2480 + 0.2945i
x_{6s}	4.0841 + 1.1182i	2.0710 + 0.0314i	1.6488 + 0.1862i	-3.3904 + 0.0460i	-2.4113 + 0.0064i
x_{7s}	4.7222 + 0.2703i	2.0603 + 0.0416i	1.7330 + 0.3341i	-3.4249 + 0.2581i	-2.4946 + 0.0075i
x_{8s}	-4.6957 - 9.5520i	4.4050 + 0.4902i	-3.9583 + 1.0601i	2.0669 + 0.0402i	1.7252 + 0.2585i
x_{9s}	4.2578 + 9.8226i	-1.9753 - 6.3435i	4.4381 + 0.4786i	2.0663 + 0.0428i	1.7416 + 0.2685i
x_{10s}	-4.9984 + 2.1066i	-0.7756 + 4.2933i	4.3961 + 0.3496i	-0.4218 - 2.6168i	1.0959 - 1.6983i

	6	7	8	9	10
x_{6s}	-1.5958 + 0.0051i	0	0	0	0
x_{7s}	-1.0196 + 1.0130i	-1.7567 - 0.0136i	0	0	0
x_{8s}	-3.1386 - 0.0071i	-2.3089 + 0.1280i	-1.5613 + 0.1695i	0	0
x_{9s}	-3.7804 + 0.4639i	-2.9044 - 0.0776i	-2.1904 + 0.0914i	-1.5003 + 0.2201i	0
x_{10s}	2.0778 + 0.0382i	1.7724 + 0.1972i	-3.3502 + 0.2027i	-2.3679 + 0.0743i	-1.5104 + 0.0398i

	1	2	3	4	5
λ_{1s}	1	0	0	0	0
λ_{2s}	0.2912 - 0.0001i	0.7088 + 0.0001i	0	0	0
λ_{3s}	0.7020 + 0.0003i	0.2349 - 0.0021i	0.0631 + 0.0018i	0	0
λ_{4s}	0.0685 - 0.0048i	0.2443 + 0.0039i	0.6925 - 0.0153i	-0.0053 + 0.0163i	0
λ_{5s}	0.9901 + 0.2637i	-0.3355 - 0.2584i	0.0828 - 0.0241i	0.0465 + 0.1288i	0.2162 - 0.1101i
λ_{6s}	0.0006 - 0.0083i	0.6419 + 0.0132i	0.0423 - 0.0125i	0.0329 + 0.0085i	0.0605 - 0.0277i
λ_{7s}	0.0025 - 0.0017i	0.7484 + 0.0257i	-0.0978 - 0.0347i	-0.0085 + 0.0152i	0.2030 - 0.1345i
λ_{8s}	-0.0000 - 0.0000i	0.0034 - 0.0030i	0.0014 - 0.0143i	0.7298 - 0.0204i	-0.0672 + 0.0228i
λ_{9s}	0.0000 + 0.0000i	-0.0001 - 0.0000i	0.0028 - 0.0031i	0.7408 - 0.0053i	-0.0809 + 0.0063i
λ_{10s}	-0.0001 - 0.0003i	-0.0004 + 0.0002i	0.0017 - 0.0008i	0.0050 - 0.0081i	0.0457 - 0.0275i

	6	7	8	9	10
λ_{6s}	0.2217 + 0.0268i	0	0	0	0
λ_{7s}	0.0466 - 0.1142i	0.1057 + 0.2442i	0	0	0
λ_{8s}	0.1429 - 0.0048i	-0.1758 + 0.1050i	0.3654 - 0.0853i	0	0
λ_{9s}	-0.0137 - 0.0107i	0.1890 - 0.0748i	-0.1868 + 0.2360i	0.3489 - 0.1482i	0
λ_{10s}	0.9435 - 0.1280i	-0.2854 + 0.2116i	0.0224 - 0.0024i	0.1463 + 0.0739i	0.1213 - 0.1186i

(j,k,k')	$x_{j,k}$	$x_{j,k'}$	$\lambda_{j,k}$	$\lambda_{j,k'}$
(2,2)	1.9896 - 0.0004i	0	0.7088 + 0.0001i	0
(3,1)	2.0250 - 0.0008i	0	0.7020 + 0.0003i	0
(4;3,4)	2.0724 + 0.0102i	1.5103 + 0.0818i	0.6925 - 0.0153i	-0.0053 + 0.0163i
(5;1,2)	2.0326 + 0.0769i	2.0005 + 0.3007i	0.9901 + 0.2637i	-0.3355 - 0.2584i
(6;2,3)	2.0710 + 0.0314i	1.6488 + 0.1862i	0.6419 + 0.0132i	0.0423 - 0.0125i
(7;2,3)	2.0603 + 0.0416i	1.7330 + 0.3341i	0.7484 + 0.0257i	-0.0978 - 0.0347i
(8;4,5)	2.0669 + 0.0402i	1.7252 + 0.2585i	0.7298 - 0.0204i	-0.0672 + 0.0228i
(9;4,5)	2.0663 + 0.0428i	1.7416 + 0.2685i	0.7408 - 0.0053i	-0.0809 + 0.0063i
(10;6,7)	2.0778 + 0.0382i	1.7724 + 0.1972i	0.9435 - 0.1280i	-0.2854 + 0.2116i

B.3 Study of the case $s = 0.465$

Now we give the zeros and the weights obtained when $s = 0.465$. By proposition 4.1.3, Moments are not the ones of a real variable. You should not thus expect that 2j first moments are the ones of a real variable. We could thus have complex zeros or weights $\lambda_{i,s}$ negative or complex.

	1	2	3	4	5
x_{1s}	0.7999 + 0.0012i	0	0	0	0
x_{2s}	-2.0974 + 0.0023i	1.9723 + 0.0019i	0	0	0
x_{3s}	1.9824 + 0.0022i	-3.4003 - 0.1706i	-1.9981 - 0.0049i	0	0
x_{4s}	2.2226 - 0.5011i	1.9998 + 0.0213i	-2.5952 - 0.0633i	-1.7760 - 0.0326i	0
x_{5s}	-0.9301 + 2.5063i	1.9726 + 0.0172i	1.5642 - 0.6725i	-2.6105 + 0.0504i	-1.7935 + 0.0750i
x_{6s}	2.1542 - 4.3900i	1.9779 + 0.0315i	1.8673 - 0.5109i	-2.0882 + 1.6927i	-2.6140 + 0.0609i
x_{7s}	-4.0445 - 1.1166i	-2.5830 - 0.0889i	-1.1113 + 1.7324i	-1.7869 - 0.0422i	1.9629 + 0.0175i
x_{8s}	7.4286 + 7.6357i	-3.8099 - 0.4467i	-2.5200 - 0.0939i	-0.5367 + 1.4318i	-1.7567 - 0.0431i
x_{9s}	5.5073 - 1.5124i	-2.1892 + 3.9845i	-3.6539 + 0.5069i	1.9638 + 0.0170i	1.4868 - 0.5080i
x_{10s}	0.7514 - 9.8259i	4.9901 + 1.0882i	-2.5534 + 2.4457i	-3.6633 + 0.6709i	1.9612 + 0.0252i

	6	7	8	9	10
x_{6s}	-1.7695 + 0.0840i	0	0	0	0
x_{7s}	1.5065 - 0.5995i	0.6621 + 0.3326i i	0	0	0
x_{8s}	1.9650 + 0.0277i	1.6981 - 0.4255i	0.0348 - 0.8821i	0	0
x_{9s}	0.0590 + 1.3782i	-0.3983 - 1.4185i	-2.4776 + 0.1129i	-1.7223 + 0.1247i	0
x_{10s}	1.6518 - 0.4035i	-0.6439 - 1.6753i	0.0349 + 1.3887i	-2.4750 + 0.1175i	-1.7020 + 0.1120i

	1	2	3	4	5
λ_{1s}	1	0	0	0	0
λ_{2s}	0.2881 + 0.0005i	0.7119 - 0.0005i	0	0	0
λ_{3s}	0.7171 - 0.0011i	0.0395 - 0.0075i	0.2434 + 0.0086i	0	0
λ_{4s}	0.0376 + 0.0754i	0.6511 - 0.0797i	0.0877 - 0.0209i	0.2236 + 0.0252i	0
λ_{5s}	0.0324 - 0.0156i	0.8350 - 0.1030i	-0.1562 + 0.0429i	0.1443 - 0.0471i	0.1445 + 0.1227i
λ_{6s}	0.0017 - 0.0023i	0.6944 - 0.2338i	-0.0369 + 0.2284i	-0.0359 - 0.0280i	0.0879 - 0.0513i
λ_{7s}	-0.0089 - 0.0072i	0.1394 + 0.1664i	-0.0738 + 0.0483i	0.0841 - 0.3429i	1.1414 - 0.2598i
λ_{8s}	0.0000 + 0.0000i	0.0167 - 0.0099i	-0.0673 - 0.0069i	0.0019 - 0.1350i	0.5480 + 0.0057i
λ_{9s}	-0.0001 + 0.0004i	0.0019 - 0.0009i	-0.0101 + 0.0227i	1.0133 - 0.2299i	-0.4804 + 0.0335i
λ_{10s}	-0.0000 - 0.0000i	0.0001 - 0.0009i	-0.0032 - 0.0145i	-0.0220 + 0.0079i	1.0615 - 0.2792i

	6	7	8	9	10
λ_{6s}	0.2888 + 0.0871i	0	0	0	0
λ_{7s}	-0.5185 - 0.2200i	0.2364 + 0.6152i	0	0	0
λ_{8s}	1.0290 - 0.4143i	-0.5501 + 0.2741i	0.0218 + 0.2863i	0	0
λ_{9s}	0.1836 - 0.0825i	-0.0874 + 0.0608i	0.1150 - 0.2117i	0.2642 + 0.4075i	0
λ_{10s}	-0.4896 + 0.1238i	-0.0622 + 0.0012	0.1594 - 0.0791ii	0.2280 - 0.2382i	0.1280 + 0.4789i

(j,k,k',k")	$x_{j,k}$	$x_{j,k'}$	$x_{j,k"}$	$\lambda_{j,k}$	$\lambda_{j,k'}$	$\lambda_{j,k"}$
(2,2)	1.9723 + 0.0019i	0	0	0.7119 - 0.0005i	0	0
(3,1)	1.9824 + 0.0022i	0	0	0.7171 - 0.0011i	0	0
(4;1,2)	2.2226 - 0.5011i	1.9998 + 0.0213i	0	0.0376 + 0.0754i	0.6511 - 0.0797i	0
(5;2,3)	1.9726 + 0.0172i	1.5642 - 0.6725i	0	0.8350 - 0.1030i	-0.1562 + 0.0429i	0
(6;1,2,3)	2.1542 - 4.3900i	1.9779 + 0.0315i	1.8673 - 0.5109i	0.0017 - 0.0023i	0.6944 - 0.2338i	-0.0369 + 0.2284i
(7;5,6)	1.9629 + 0.0175i	1.5065 - 0.5995i	0	1.1414 - 0.2598i	-0.5185 - 0.2200i	0
(8;6,7)	1.9650 + 0.0277i	1.6981 - 0.4255i	0	1.0290 - 0.4143i	-0.5501 + 0.2741i	0
(9;4,5)	1.9638 + 0.0170i	1.4868 - 0.5080i	0	1.0133 - 0.2299i	-0.4804 + 0.0335i	0
(10;5,6)	1.9612 + 0.0252i	1.6518 - 0.4035i	0	1.0615 - 0.2792i	-0.4896 + 0.1238i	0

B.4 Conclusion

In the example studied here, we see that it is thus possible to detect zeros easily and also value of s. For example, if we choose s=0,465, we see weights bigger than 1 : they indicate that we have exceeded the possible values in order to have a sequence of not imaginary real variables. It is combined with negative weights. We can thus look from when we have too important negative weights to try to determine s. By studying these points for various values of s, we shall thus have curves as in section 4.2, what will allow us to determine easily s

In fact this behavior can depend on samples or on parameters: it will be maybe not the same thing for other samples or other mixtures. It will thus be necessary to make a more complete study of the behavior of the weights $\lambda_{j,s}$ when s increases.

However in examples as above, we can determine easily the first s, the zeros and the variance.

Let us remark that one could also estimate the moments of Y+isU without using samples of Y+isU, but by computing real moments of the Gaussian random variable U :

$$\mathbb{E}_e\{(Y+isU)^j\} = \sum_{k=0}^{j} \frac{j!}{k!(j-k)!}\mathbb{E}_e\{Y^k\}\mathbb{E}\{(isU)^{k-j}\} = \sum_{k=0}^{j} \frac{j!}{k!(j-k)!}\mathbb{E}_e\{Y^k\}(i.s)^{k-j}\mathbb{E}\{U^{k-j}\} \ ,$$

where the moments $\mathbb{E}\{U^{k-j}\}$ are the Gaussian moments of N(0,1) and where $\mathbb{E}_e\{Y^k\}$ are empirical moments of the sample of Y.

Let us notice now that this example is an example easy to study. Indeed in that case, the smallest variance is associated with the biggest weight. It is thus normal that the zero $s^2 = 2$ show up fast. Otherwise, it could be more delicate. It will then be necessary to take values of j bigger. It will thus depend on computers and on necessary programs to calculate the orthogonal polynomials. But with the advances of the researches in these domains, we can hope to improve this kind of result .

As deja seen in [9], we can finally hope to have simple methods to detect the components of Gaussian mixtures.

But here, it is not the purpose of this report : this purpose, it is to show some possibilities of the variances of higher order and of the quadrature of Gauss Jacobi in Probability and Statistics. A more complete study can be done later in order to understand if we can really obtain programs which allow to find the components of a Gaussian mixture with efficiency and can be even more efficiency than the algorithm EM.

Bibliography

[1] LANCASTER H. O. (1960) Orthogonal models for contingency tables. Developments in statistics. Academic Press, New York.

[2] BLACHER R. (1993) Higher Order Correlation Coefficients. Statistics 25, 1-15.

[3] BLACHER R. (2012) Higher Order Correlation Coefficients. Hal Archives ouvertes http://hal.archives-ouvertes.fr/hal-00734568

[4] BLACHER R. (1995) Central limit theorem by polynomial dependence coefficients. Journal of computational and applied mathematics 57, 45-56.

[5] BLACHER R. (1983) Indicateurs de dépendance fournis par le développement en série de la densité de probabilité. Thèse 3 Cycle, Université Joseph Fourier de Grenoble. Hal archives ouvertes : http://tel.archives-ouvertes.fr

[6] BLACHER R. (1990) Quelques applications des fonctions orthogonales en probabilités et statistiques, Thèse de l'université Joseph Fourier, Grenoble 1. Hal archives ouvertes : http://tel.archives-ouvertes.fr

[7] BLACHER R. (1984) Coefficient de correlation d'ordre (i,j) et variances d'ordre i. Rapport de recherche 463, Laboratoire TIM3 IMAG, Université de Grenoble.

[8] BLACHER R. (1984) Coefficient de correlation d'ordre (i,j) et variances d'ordre i. Symposium International de Bar Le Duc, Polynomes orthogonaux et applications, Springer Verlag, 1171, 475-486.

[9] BLACHER R. (2012) Higher Order variance and Gauss Jacobi Quadrature http://hal.archives-ouvertes.fr/hal-00708152/

[10] BLACHER R. (2003) Multivariate quadratic forms of random vectors. Journal of multivariate analysis 87, 2-23.

[11] MCLACHLAN G.J. KRISHNAN T. (1997) The EM algorithm and extensions, second edition, Wiley series in probability and statistics. Wiley .

[12] SERFLING R.J. (1980) Approximation theorems of mathematical statistics. Wiley, New York

[13] CHIHARA. (1978) An introduction to orthogonal polynomials Gordon and Breach, New York

[14] FREUD Geza. (1971) Orthogonal polynomials. Pergamon Press

[15] HIDA T. (1980) Brownian motion. Springer, Berlin

[16] SZEGO G. (1959) Orthogonal polynomials. Americal Mathematical Society, Providence.