

Brique BASTA

Étude de cas :

Étude de la distribution des sauts d'amplitude dans une image
numérique

Audran LE BARON

17 décembre 2002



FIG. 1 – Image de travail

Les données

Nous travaillerons tout au long de cette étude de cas sur l'image `aero.tiff`, image en 256 niveaux de gris de taille 510×341 . Nous chargeons cette image et l'affichons par le code MatLab suivant. L'image est représentée figure 1.

```
a=imread('aero.tiff');  
imagesc(a);  
colormap gray;  
colorbar;  
axis equal;
```

Nous travaillerons plus particulièrement sur le vecteur des sauts d'amplitude horizontaux, noté x , obtenu par le code suivant :

```
a = double(a);  
[M, N] = size(a);  
h = a(:,2:end) - a(:,1:end-1);    % Matrice des sauts horizontaux  
                                     % (taille M x N-1)  
x = reshape(h, prod(size(h)), 1);  % Matrice aplatie en vecteur
```

La figure 2 représente les sauts d'amplitude horizontaux de cette image.

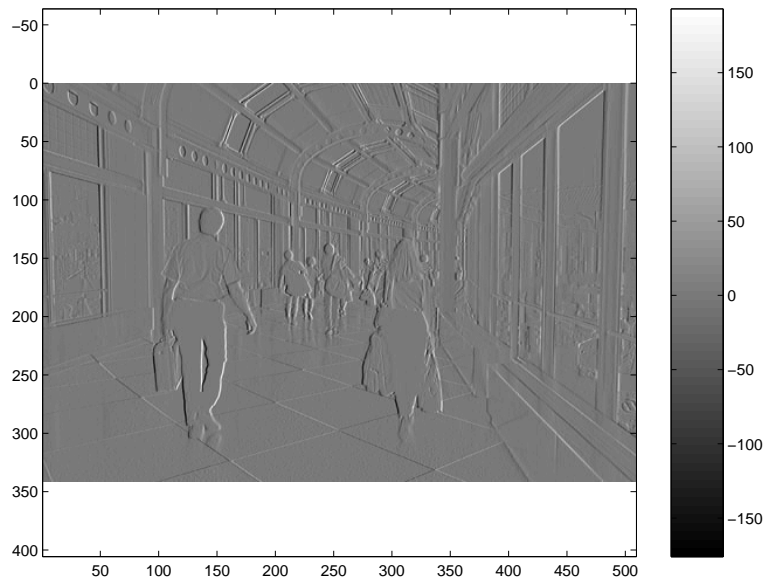


FIG. 2 – Sauts d’amplitude horizontaux

1 Modélisation par une gaussienne

Considérons pour commencer le modèle paramétrique dans lequel la densité de notre variable aléatoire (valeur du saut d’amplitude) suit une loi gaussienne :

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

1.1 Comparaison d’histogrammes

Le calcul de la moyenne et de la variance des observations se fait par le code suivant :

```
m = mean(x);
v = var(x);
```

Les résultats sont :

$$\begin{aligned}\mu &= 0,0157 \\ \hat{\sigma}^2 &= 337,3939\end{aligned}$$

Par le code suivant, on trace la distribution gaussienne dont les paramètres sont les moyenne et variance empiriques, avec l’histogramme des observations. Le résultat n’est pas très satisfaisant comme le montre la figure ??.

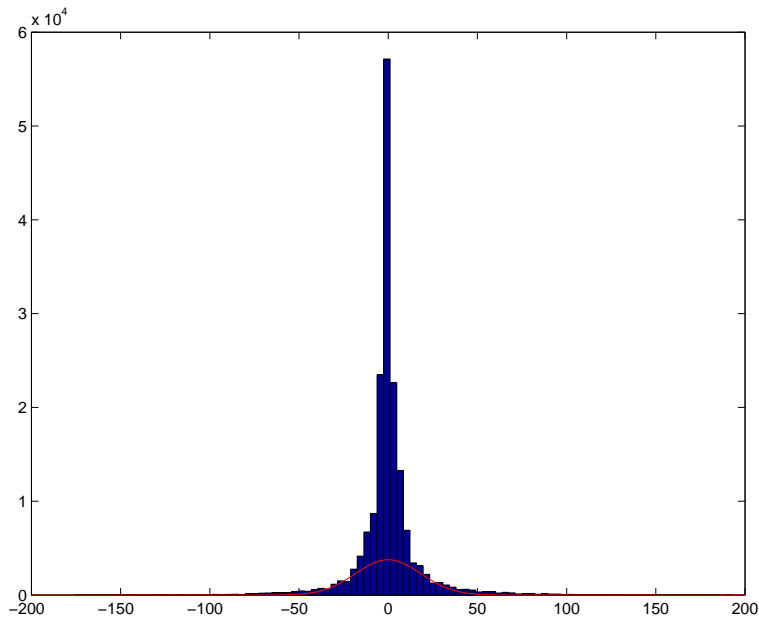


FIG. 3 – Histogramme des observations et distribution gaussienne

```
x_range = [-200:.1:200];
gauss_m_v = 1/(sqrt(2*pi*v)) * exp(-(x_range-m).^2 ./ (2*v));

hist(x, 100);
hold on
plot(x_range, length(x) * gauss_m_v, 'r');
hold off;
```

1.2 Comparaison de quantiles

Soit $F_{\mu,\sigma}$ la fonction de répartition d'une distribution gaussienne de paramètres (μ, σ) . Nous noterons $F = F_{1,1}$. Le quantile d'ordre p de $g_{\mu,\sigma}$ est le x tel que :

$$\begin{aligned}
 p = F_{\mu,\sigma}(x) &= \int_{-\infty}^x \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt \\
 &= \frac{1}{\sigma} \int_{-\infty}^{\frac{x-\mu}{\sigma}} \exp\left(-\frac{t^2}{2}\right) dt \\
 &= \frac{1}{\sigma} F\left(\frac{x-\mu}{\sigma}\right)
 \end{aligned}$$

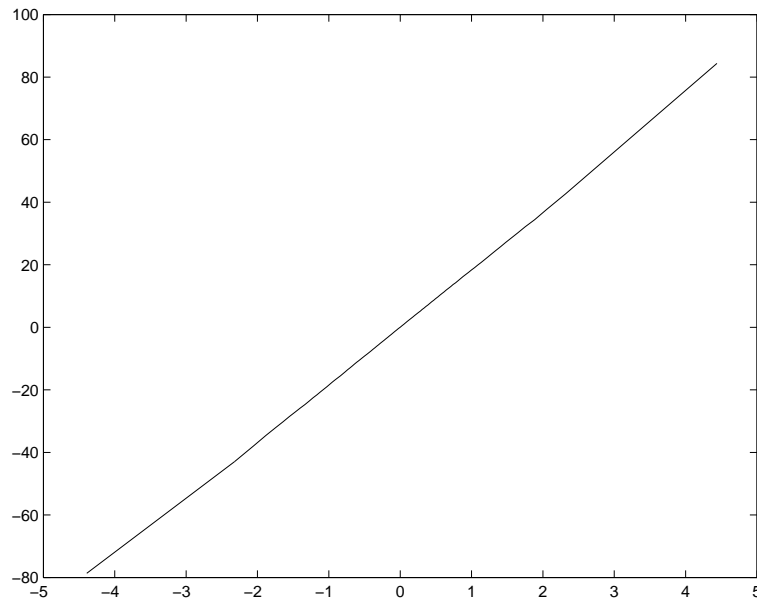


FIG. 4 – Quantiles d’une loi gaussienne en fonction des quantiles d’une autre loi gaussienne

Ainsi, si $q_{\mu,\sigma}(p)$ est le quantile d’ordre p d’une gaussienne (μ, σ) et $q(p)$ celui d’une gaussienne centrée réduite, on a :

$$q_{\mu,\sigma}(p) = \sigma q(p) + \mu$$

C’est pourquoi, lorsque l’on trace les quantiles d’une variable gaussienne par rapport aux quantiles d’une seconde variable gaussienne, on obtient une droite (Cf. figure 4). Ce n’est pas le cas lorsque l’on trace les quantiles des observations (Cf. figure 5) en fonction des quantiles d’une loi normale par le code suivant, comme le montre la figure 6.

```
g_m_v = m + sqrt(v) * randn(length(x), 1);
qqnorm(g1); % C’est bien une droite

qqnorm(x); % Ce n’est PAS une droite
```

On note qu’il y a trop de poids dans la queue de la gaussienne et pas assez au « centre ».

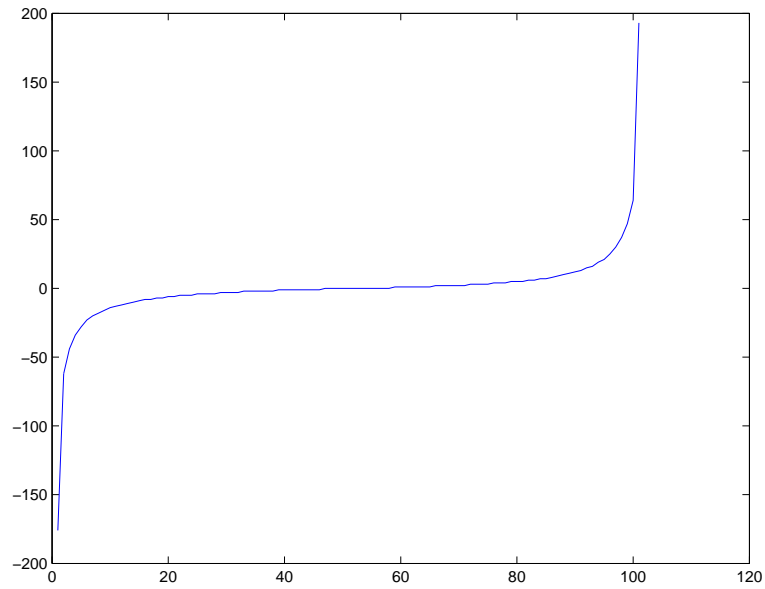


FIG. 5 – Quantiles des observations

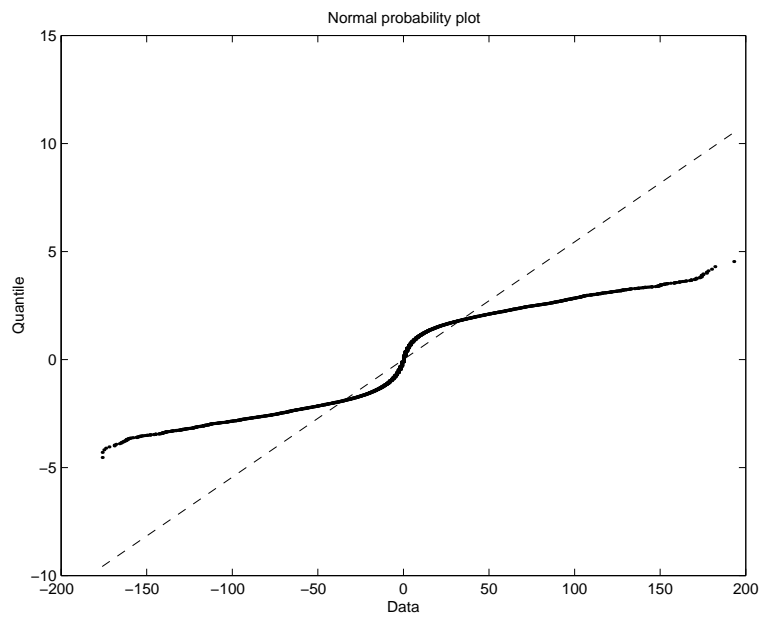


FIG. 6 – Quantiles des observations en fonction des quantiles d’une loi normale

2 Modélisation par une loi « laplacienne généralisée »

Changeons de modèle paramétrique pour prendre celui d'une loi de Laplace généralisée, de densité :

$$h(x) = \frac{\alpha\eta}{2\Gamma(1/\alpha)} \exp(-|\eta x|^\alpha)$$

2.1 Estimation par la méthode des moments

Une première méthode d'estimation est la méthode des moments. Nous avons deux paramètres scalaires (α et η), donc deux moments qui sont ici la variance et le kurtosis, donnés par :

$$\sigma^2 = \frac{\Gamma\left(\frac{3}{\alpha}\right)}{\eta^2\Gamma\left(\frac{1}{\alpha}\right)}$$
$$\kappa = \frac{\Gamma\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{5}{\alpha}\right)}{\Gamma\left(\frac{3}{\alpha}\right)^2}$$

Le code suivant permet de tracer $\kappa(\alpha)$ et de calculer le kurtosis empirique $\hat{\kappa}$.

```
alpha_vect = [25:100]/50;  
k = kurt(alpha_vect);  
plot(alpha_vect, k);
```

```
k_emp = kurtosis(x);
```

$\kappa(\alpha)$ est représenté figure 7 et le calcul du kurtosis empirique donne $\hat{\kappa} = 17,8968$. Donc d'après la courbe, α se trouverait entre 0,5 et 0,6.

D'où la résolution numérique suivante :

```
alpha_vect = [0.5:.0001:0.6];  
[trash, indice] = min(abs(kurt(alpha_vect) - k_emp))  
alpha = alpha_vect(indice)  
eta = 1 / sqrt(v * gamma(1/alpha) / gamma(3/alpha))
```

Le résultat de ce calcul donne :

$$\alpha = 0,5672$$
$$\eta = 0,3469$$

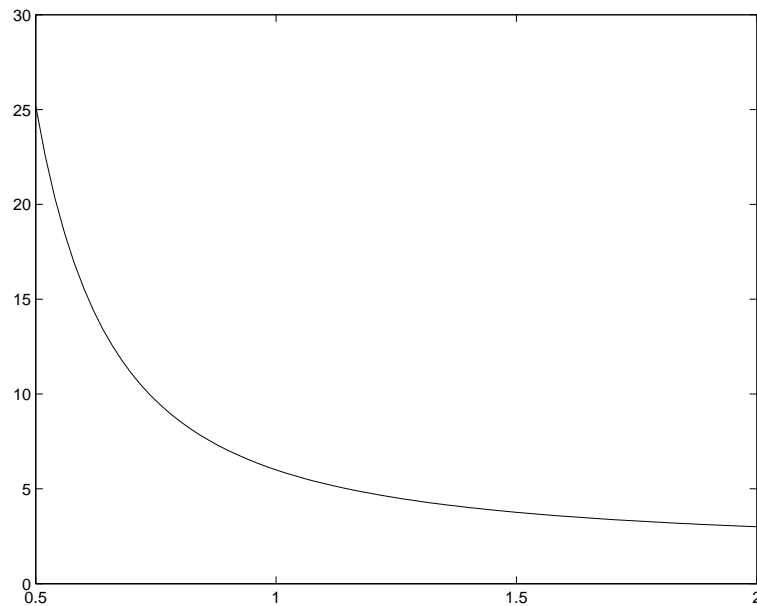


FIG. 7 – Représentation de $\kappa(\alpha)$

Le code suivant fait représenter sur le même graphique (*Cf.* figure 8) l'histogramme des observations, la distribution gaussienne trouvée à la section précédente et la distribution de Laplace généralisée avec les paramètres estimés par la méthode des moments.

```
x_range = [-200:.1:200];
densite_laplace = alpha * eta / (2 * gamma(1/alpha)) * exp(-abs(eta * x_range).^alpha);
gauss_m_v = 1/(sqrt(2*pi*v)) * exp(-(x_range-m).^2 ./ (2*v));
hist(x, 100);
hold on
plot(x_range, length(x) * densite_laplace, 'r')
plot(x_range, length(x) * gauss_m_v, 'g');
hold off;
```

Afin de mieux comparer, on trace les quantiles de cette loi de Laplace en fonction des quantiles empiriques des observations. Cela demande un calcul préalable des approximations des quantiles de la loi.

```
x_range = [min(x):0.1:max(x)];
repartition = 1/2 * (1 + sign(x_range)) .* gammainc(abs(eta .* x_range) .^ alpha);
p_range = [0.01:0.01:1];
for i = 1:100
```

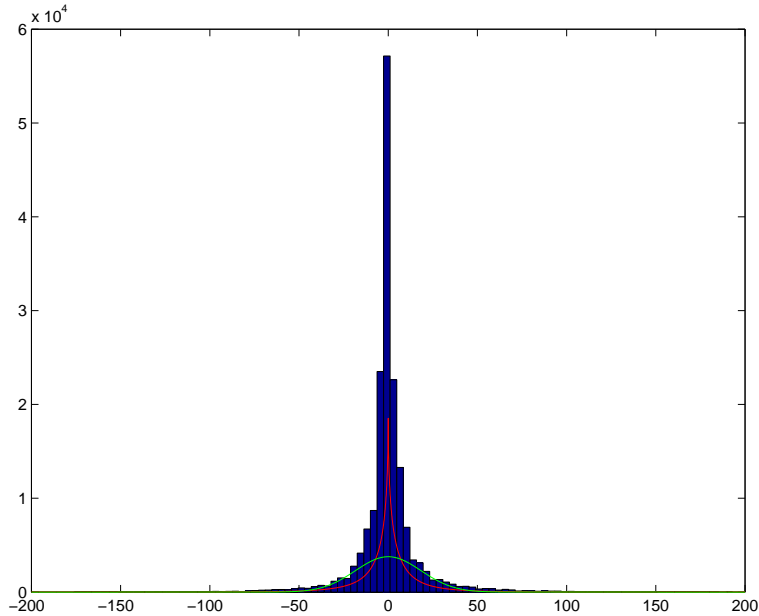


FIG. 8 – Observations, gaussienne et Laplace (moments)

```

[temp, indice] = min(abs(repartition - p_range(i)));
quant_emp(i) = x_range(indice);
end
plot(quantile(x, p_range), quant_emp);

```

On constate une certaine amélioration par rapport au modèle gaussien (voir figure 9). Il y a cependant toujours trop de poids dans la queue de la distribution de Laplace.

2.2 Estimation par la méthode du maximum de vraisemblance

Une autre méthode d'estimation des paramètres est la méthode du maximum de vraisemblance. Ici, les deux équations de vraisemblance sont :

$$1/\alpha + \Psi(1/\alpha)/\alpha^2 - \frac{1}{n} \sum_{i=1}^n \log(n|X_i|) |X_i|^\alpha$$

$$\eta = \left(\frac{\alpha}{n} \sum_{i=1}^n |X_i|^\alpha \right)^{-\frac{1}{\alpha}}$$

Par itérations successives et grâce au code suivant, on calcule $\tilde{\alpha}$ et $\tilde{\eta}$.

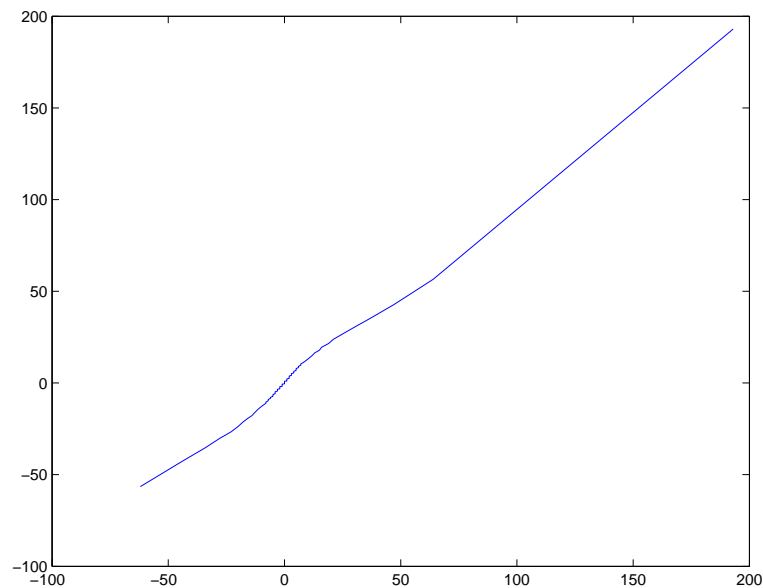


FIG. 9 – Quantiles de la loi de Laplace en fonction des quantiles empiriques (moments)

```

% Suppression des zeros
[trash1 trash2 y] = find(x);

% Initialisation
n = length(x);

alpha = 0.5672;
eta = 0.3469;

pas = 0.01;
ecart = 0.5;

logvrais = -Inf;

% Boucle
alpha_vect = [alpha - ecart : pas : alpha + ecart];
for i=1:30
    [trash indice] = min(abs(eq4(y, alpha_vect, eta, n)));
    alpha = alpha_vect(indice);
    eta = power(alpha / n * sum (power(abs(x), alpha)), -1/alpha);
    newlogvrais = sum(log(laplace(x, alpha, eta)));

```

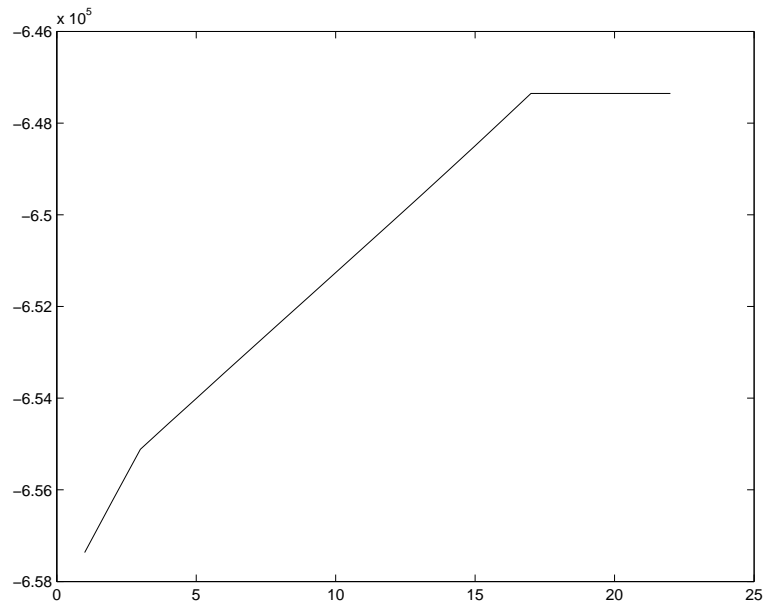


FIG. 10 – Évolution de la log-vraisemblance à chaque itération

```

i
alpha
eta
newlogvrais
(newlogvrais - logvrais)/abs(newlogvrais)
logvrais = newlogvrais;
logvraisvect(i) = logvrais;
end ;
plot(logvraisvect);

```

L'évolution de la log-vraisemblance au fur et à mesure des itérations est représenté figure 10. Au bout de 18 itérations, la log-vraisemblance se stabilise pour les valeurs :

$$\tilde{\alpha} = 0,4072$$

$$\tilde{\eta} = 1,7715$$

On peut alors, pour ces nouvelles valeurs des paramètres, retracer la distribution et les quantiles. Ceci est fait aux figures 11, 12. On constate à nouveau une amélioration, mais le modèle n'est toujours pas parfait.

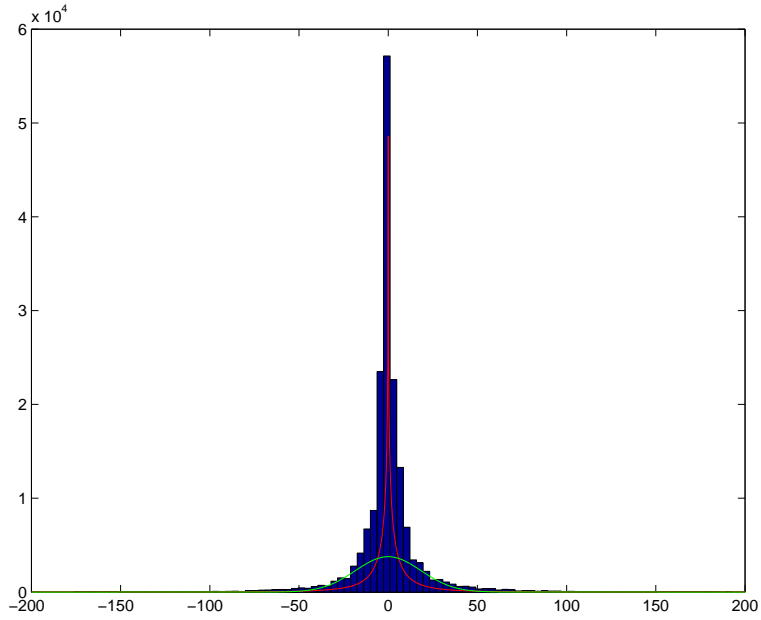


FIG. 11 – Observations, gaussienne et Laplace (max. de vraisemblance)

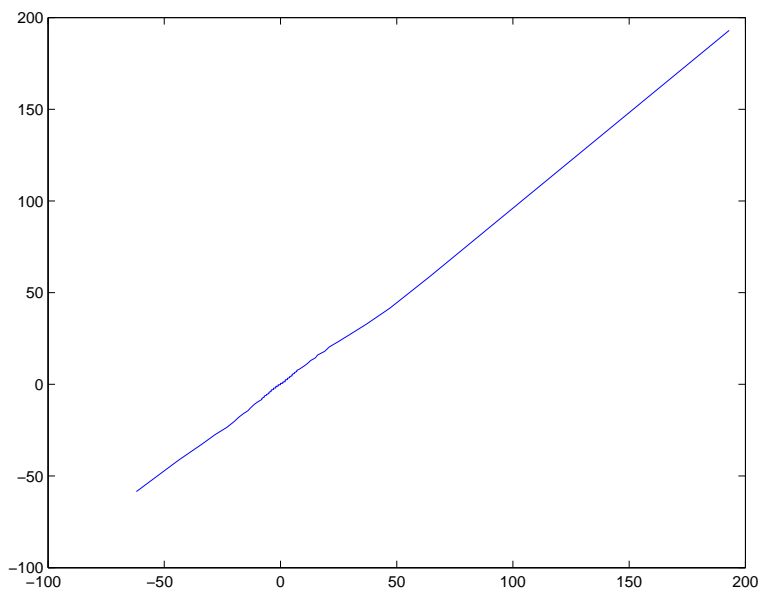


FIG. 12 – Quantiles de la loi de Laplace en fonction des quantiles empiriques des observations (max.de vraisemblance)

3 Modélisation par un mélange de gaussiennes

Un autre modèle possible est celui du mélange de gaussiennes :

$$f_{\theta}(x) = \sum_{k=1}^K \alpha_k g_{\theta_k}(x)$$

Un début d'implémentation a été effectué, mais encore aucun résultat n'a pu être obtenu, faute de temps...