TP 1 — Segmentation bayésienne

Probabilistic Models and Machine Learning

Ambroise LAROYE–LANGOUËT

Télécom Paris

11 octobre 2025

1 Introduction

Ce TP a pour objectif de mettre en œuvre différentes méthodes de segmentation de signaux, notamment avec le **Maximum de Vraisemblance** et le **Maximum A Posteriori**.

2 Segmentation par Maximum de Vraisemblance (MV)

2.1 Principe

La segmentation par Maximum de Vraisemblance consiste à associer chaque observation y_s à la classe ω_i maximisant la densité de probabilité $p(y_s|\omega_i)$.

$$\hat{s}(y) = \arg\max_{\omega_i} p(y|\omega_i)$$

2.2 Implémentation

— `bruit_gauss(X, cl1, cl2, m1, sig1, m2, sig2)`
— `classif_gauss2(Y, cl1, cl2, m1, sig1, m2, sig2)`
— `taux_erreur(A, B)`

Nous utilisons norm.pdf directement, pour implémenter la fonction `classif_gauss2`, l'objectif est de classer chaque élément du signal bruité dans l'une des deux classes (cl1 ou cl2) en utilisant le critère du maximum de vraisemblance. Cela signifie que pour chaque valeur, on calcule la probabilité qu'elle provienne de la gaussienne associée à cl1 ou à cl2, et on choisit la classe la plus probable.

La densité de probabilité d'une gaussienne est :

$$P(y|m,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right)$$

Selon la classe Cl1 ou Cl2 nous comparons les densités de probabilité et nous associons la plus probable à chaque fois.

2.3 Résultats

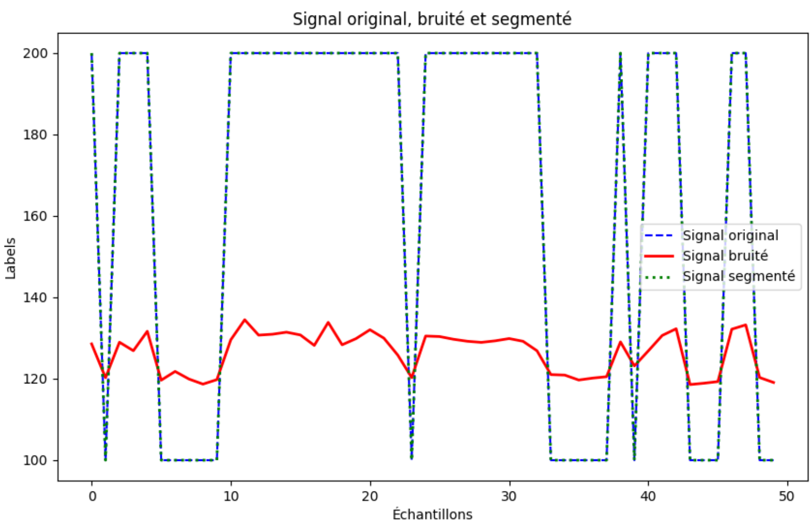


Figure 1 – Signal original, bruité et segmenté (méthode MV).

Taux d'erreur de segmentation **0.0%**.
Nous observons que le signal original et le signal segmenté sont superposés, ce qui suggère que notre méthode à bien fonctionné.

2.4 Taux d'erreur moyen

Afin de mesurer statistiquement l'erreur, il est nécessaire de moyenner celle-ci sur un grand nombre T de versions bruitées du même signal, simulées à paramètres constants.
T est le nombre de fois où nous bruitons le signal original avec les mêmes paramètres de bruit et nous classons le signal bruité obtenu pour estimer X_{est}.
De ce fait nous avons unee stimation plus robuste. En effet, une seule simulation peut donner une erreur très variable, car le bruit est aléatoire. En répétant l'expérience T fois, nous obtenons une moyenne des erreurs qui est plus représentative de la performance de cette méthode de classification.

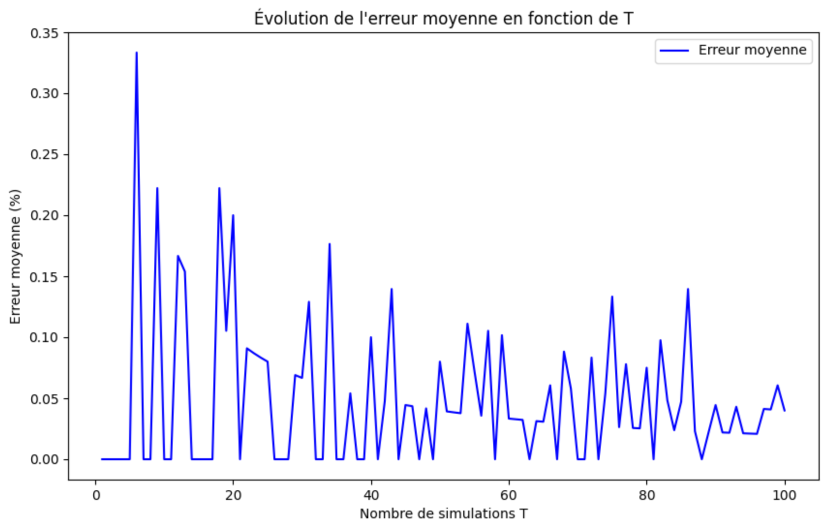


FIGURE 2 – Graph de l'évolution de l'erreur moyenne selon T

Nous remarquons que l'erreur **converge** se rapprochant d'une valeur autour de 0.05. Plus T est grand, plus la moyenne des erreurs se rapproche de la "vraie" erreur moyenne obtenue avec un nombre infini de simulations (loi des grands nombres).
Pour que ce soit plus visible nous pourrions augmenter T, et visualiser pour T=300 par exemple.

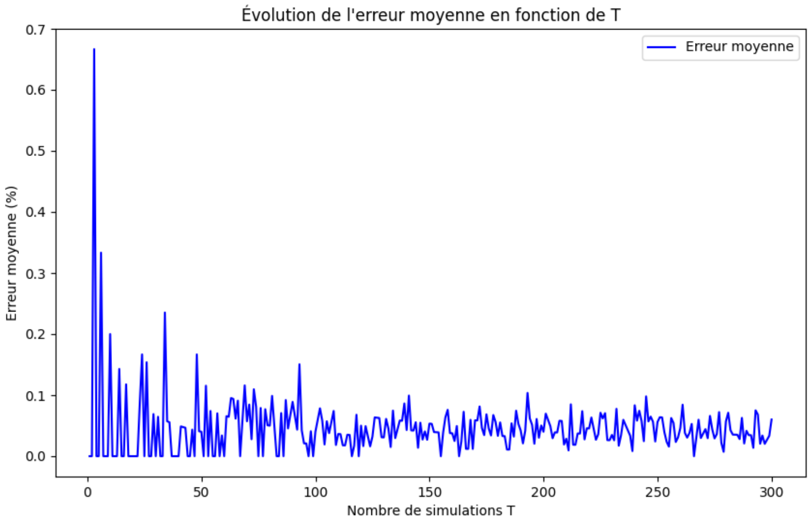


FIGURE 3 – Graph de l'évolution de l'erreur moyenne avec T=300

L'erreur moyenne est faible et stable à partir de T = 100. Et au delà, elle prend ses valeurs autour de 0.05 et encadré par $0; 0.1$. Le bruit est faible suggérant que la classification est plutôt robuste.

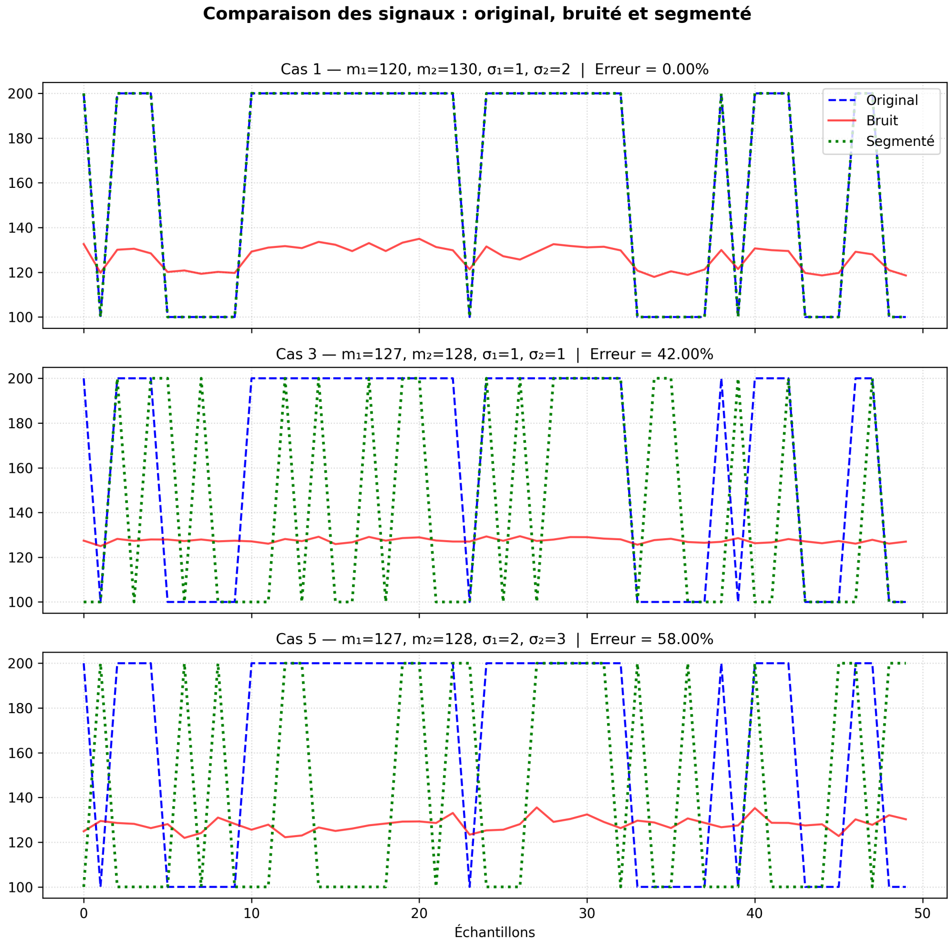
Test avec les 6 signaux

	m_1	m_2	σ_1	σ_2	**Signal 1**	**Signal 2**	**Signal 3**	**Signal 4**	**Signal 5**	**Signal 6**
1	120	130	1	2	0.04	0.03	0.04	0.04	0.04	0.04
2	127	127	1	5	20.42	17.81	17.64	17.68	17.68	17.62
3	127	128	1	1	30.54	31.28	30.86	30.83	30.86	30.87
4	127	128	0.1	0.1	0	0	0	0	0	0
5	127	128	2	3	43.28	38.42	38.07	38.21	38.21	38.09

TABLE 1 – Résultats de segmentation MV pour différents niveaux de bruit.

Cas 1 : erreurs très faibles, ce qui était attendu car peu de bruit.
Cas 2–3 : erreurs plutôt élevé, les moyennes sont trop proche et dans le cas 3 les variances sont trop élevé et il est ainsi plus difficile de distinguer les signaux.
Cas 4 : erreurs nulles. C'est le cas idéal où les variances sont très fable et les moyennes suffisamment espacées.
Cas 5 : erreurs très élevées (> 38). La variance est trop grande, le bruit est fort et donc la confusion est élevée.

Nous pouvons visualiser comme dans la figure 1, les signaux originaux, segmenté et le bruit. Afin de mieux comprendre les erreurs.


Comparaison des signaux : original, bruité et segmenté
Cas 1 — m₁=120, m₂=130, σ₁=1, σ₂=2 | Erreur = 0.00%
Original
Bruit
Segmenté
Cas 3 — m₁=127, m₂=128, σ₁=1, σ₂=1 | Erreur = 42.00%
Cas 5 — m₁=127, m₂=128, σ₁=2, σ₂=3 | Erreur = 58.00%
200
180
160
140
120
100
0
10
20
30
40
50
Échantillons

3 Méthode bayésienne : Maximum a Posteriori (MAP)

3.1 Principe

Précédemment nous avons vu la méthode du MV, c'est à dire que l'on cherchait la classe qui rend l'observation la plus probable. On cherchait à maximiser la probabilité d'observer les données, sans tenir compte des probabilités a priori des classes. i.e. on ne prenait pas en compte la fréquence d'apparition des classes dans l'ensemble.

Maintenant, nous introduisons les probabilités a priori $p(\omega_i)$:

$$\hat{s}_B(y) = \begin{cases} \omega_1 & \text{si } p(\omega_1)p(y|\omega_1) \geq p(\omega_2)p(y|\omega_2) \\ \omega_2 & \text{sinon.} \end{cases}$$

$$\hat{\omega}_{MAP} = \arg\max_{\omega_i} p(\omega_i|y)$$

En utilisant le théorème de Bayes :

$$p(\omega_i|y) = \frac{p(y|\omega_i)p(\omega_i)}{p(y)}$$

Puisque la probabilité marginale $p(y)$ est constante pour toutes les classes :

$$\hat{\omega}_{MAP} = \arg\max_{\omega_i} p(y|\omega_i)p(\omega_i)$$

Ainsi le MAP réduit les erreurs globales.

Remarque : Dans le cas indépendant, MPM = MAP. Donc ici, parler de MAP ou MPM revient au même.

3.2 Implémentation

(cf code)

- `calc_probaprio2(X, cl1, cl2)`
- `MAP_MPM2(Y, cl1, cl2, p1, p2, m1, sig1, m2, sig2)`

3.3 Résultats et Comparaisons

	m_1	m_2	σ_1	σ_2	**Signal 1**	**Signal 2**	**Signal 3**	**Signal 4**	**Signal 5**	**Signal 6**
1	120	130	1	2	0.06	0.04	0.04	0.04	0.04	0.04
2	127	127	1	5	19.90	17.91	17.66	17.71	17.69	17.62
3	127	128	1	1	27.50	30.85	30.83	30.88	30.85	30.83
4	127	128	0.1	0.1	0	0	0	0	0	0
5	127	128	2	3	36.00	38.38	38.11	38.22	38.16	38.06

TABLE 2 – Erreurs moyennes pour chaque signal selon les paramètres de bruit (méthode du Maximum A Posteriori)

1. Comme pour la méthode MV, le faible niveau de bruit et la bonne séparation des moyennes permettent une classification quasi parfaite. Le MAP apporte un léger gain.
2. Les deux classes ont la **même moyenne**, ce qui rend la séparation difficile. L'ajout des probabilités a priori ne change pas significativement les résultats. Les erreurs sont comparables à celles du MV.

3. Les moyennes sont très proches, donc la confusion reste forte. Cependant, on note une légère amélioration par rapport à la MV.
4. Comme pour MV, bruit très faible et classes bien séparées. MAP et MV donnent des performances identiques parfaites.
5. Grandes variances induisent une forte confusion. On remarque une réduction de l'erreur moyenne par rapport au MV.

De manière générale, nous remarquons des légères améliorations avec la méthode MAP par rapport à la méthode MV.

p1/p2	120/130 1/2	127/127 1/5	127/128 1/1	127/128 0.1/0.1	127/128 2/3
0.1/0.9	0.06	26.68	31.70	0.00	53.32
0.3/0.7	0.02	22.28	30.44	0.00	45.24
0.5/0.5	0.00	17.84	30.06	0.00	38.84
0.7/0.3	0.02	12.98	31.26	0.00	32.10
0.9/0.1	0.02	8.70	30.84	0.00	24.56

Table 3 – Erreurs moyennes (%) pour le Maximum de Vraisemblance (MV)

p1/p2	120/130 1/2	127/127 1/5	127/128 1/1	127/128 0.1/0.1	127/128 2/3
0.1/0.9	0.02	9.46	9.76	0.00	9.72
0.3/0.7	0.02	20.04	26.10	0.00	29.62
0.5/0.5	0.00	17.84	30.06	0.00	38.84
0.7/0.3	0.00	12.00	24.62	0.00	26.56
0.9/0.1	0.00	4.56	9.90	0.00	9.96

Table 4 – Erreurs moyennes (%) pour le Maximum A Posteriori (MAP)

Paramètres de bruit	**Résultats MV et MAP**	**Interprétation**
120/130 1/2	Les erreurs sont très faibles pour les deux méthodes.	Les moyennes des classes sont bien séparées, ce qui facilite la classification.
127/127 1/5	Les erreurs sont significativement plus faibles pour MAP que pour MV.	Les moyennes des classes sont identiques, ce qui rend la classification difficile pour le MV.
127/128 1/1	MV : Les erreurs sont élevées (environ 30%). MAP : Les erreurs sont plus faibles pour les cas où $p1$ et $p2$ sont déséquilibrés ($p1/p2 = 0.1/0.9$).	Les moyennes des classes sont proches, ce qui rend la classification difficile.
127/128 0.1/0.1	MV et MAP : Les erreurs sont nulles.	Le bruit est très faible, ce qui permet une classification parfaite pour les deux méthodes.
127/128 2/3	MV : Les erreurs sont très élevées. MAP : Les erreurs sont significativement plus faibles ($p1/p2 = 0.1/0.9$ resp $0.9/0.1$).	Le bruit est élevé et les moyennes sont proches, ce qui rend la classification difficile.

Pour conclure,

- Le MAP améliore la segmentation lorsque les probabilités a priori sont déséquilibrées.
- De plus, la méthode MAP est plus robuste face au bruit élevé et aux moyennes proches des classes.
- Le taux d'erreur diminue avec l'augmentation du rapport signal/bruit.
- Le MAP est plus performant lorsque les classes ont des fréquences différentes ou lorsque le bruit est élevé.
- Le MV est suffisant lorsque les classes sont équilibrées et le bruit est faible.
- Le MV est plus simple à implémenter et peut suffire dans certains cas.