

# Approche multi-échelles dans les espaces de formes

Raphaël Deswarte, Keurcien Luu, Justin Trias

30 juin 2011

- 1 Introduction
- 2 Méthodes linéaires
  - Appariement exact ( "exact matching" )
  - Appariement inexact ( "inexact matching" )
- 3 Méthode non linéaire
  - Approche multi-échelles
- 4 Moyenne de Karcher
- 5 Application : étude d'une maladie génétique chez les souris.
  - Cadre de l'étude
  - Analyse statistique des résultats
- 6 Conclusion

# Plan

- 1 Introduction
- 2 Méthodes linéaires
- 3 Méthode non linéaire
- 4 Moyenne de Karcher
- 5 Application : étude d'une maladie génétique chez les souris.
- 6 Conclusion

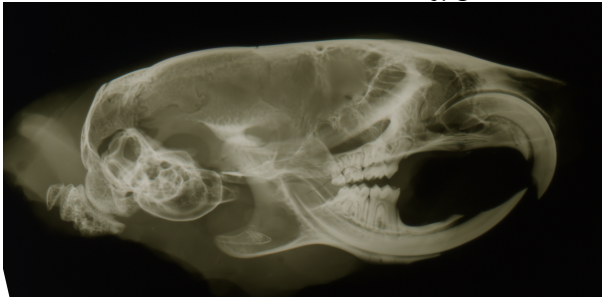
Une mutation du gène *msx2* entraîne des modifications du pelage, du comportement, des formes de la mâchoire et du crâne chez la souris.

souris mutante.jpg



Nous allons comparer les formes du crâne et de la mâchoire des souris.

male msx2 +/- 3,5 mois.jpg



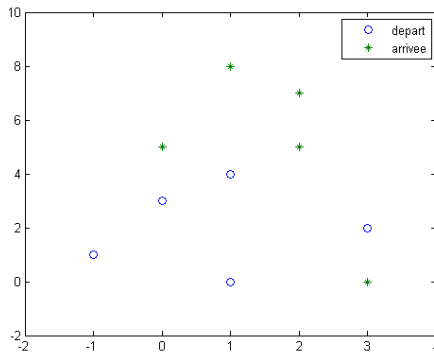
**FIGURE:** Radiographie d'un individu hétérozygote.

# Plan

- 1 Introduction
- 2 Méthodes linéaires
  - Appariement exact ( "exact matching" )
  - Appariement inexact ( "inexact matching" )
- 3 Méthode non linéaire
- 4 Moyenne de Karcher
- 5 Application : étude d'une maladie génétique chez les souris.

# Appariement exact

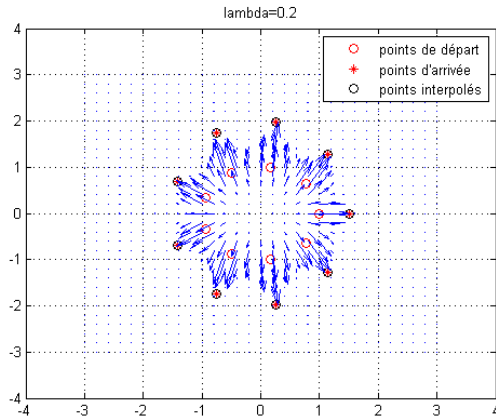
- On considère  $N$  points initiaux et  $N$  points cibles ("landmarks").

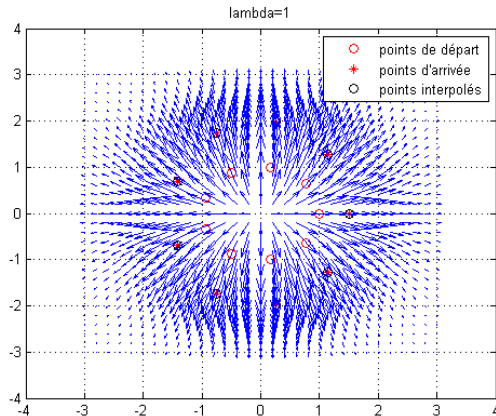


- **Théorème de Moore-Aronszajn** : Si  $K$  est un noyau défini positif sur  $X$ , alors il existe un espace de Hilbert de fonctions définies sur  $X$  tel que  $K$  est le noyau reproduisant de cet espace.

- Nous travaillons avec l'espace de Hilbert associé au noyau gaussien :  $K_V(x, y) = \exp(-\frac{|x - y|^2}{2\lambda^2}) I_2$ .
- But : déterminer le champ de déplacement  $v(y)$  de norme minimale dans cet espace.

$$\forall i, v(x_i) = y_i - x_i, 1 \leq i \leq n.$$



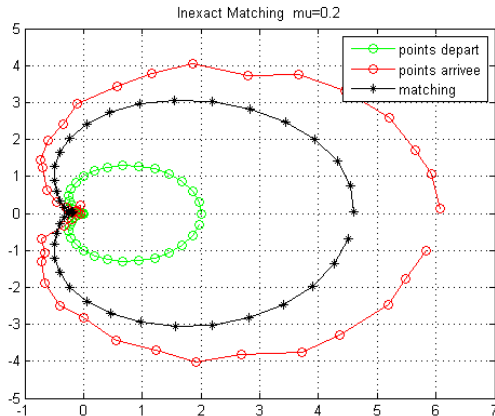


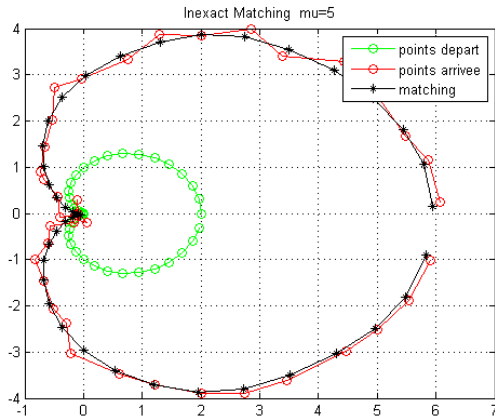
# Appariement inexact

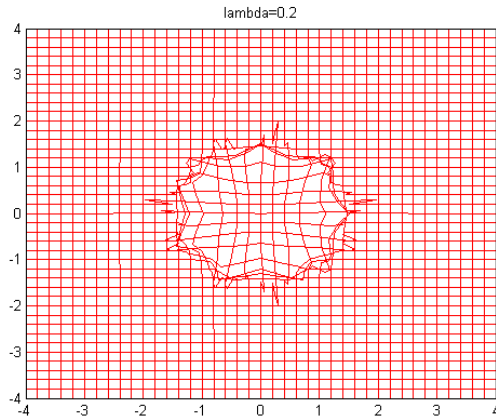
On cherche à minimiser la quantité :

$$\frac{1}{2} \|v\|_V^2 + \mu \sum_{i=1}^N |y_i - (x_i + v(x_i))|^2$$

.







# Plan

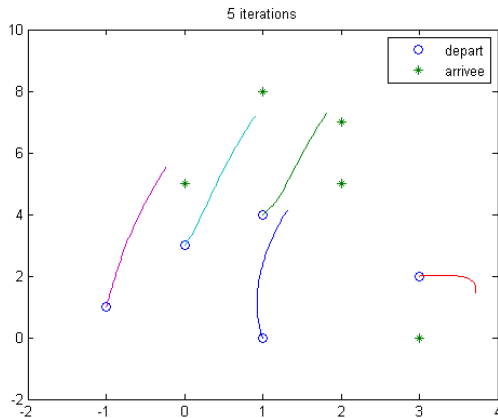
- 1 Introduction
- 2 Méthodes linéaires
- 3 Méthode non linéaire
  - Approche multi-échelles
- 4 Moyenne de Karcher
- 5 Application : étude d'une maladie génétique chez les souris.
- 6 Conclusion

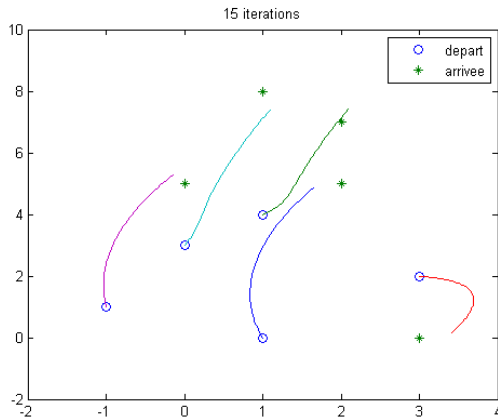
- $N$  points de départ et  $N$  points d'arrivée correspondants
- $v_x(t)(y)$  n'est plus un déplacement mais la vitesse du point  $y$  à l'instant  $t \in [0, 1]$
- Minimiser la fonctionnelle  $J$  par rapport à  $\alpha$  sous la contrainte  $\dot{x} = S_V(x) \alpha$  et  $x(0)$  fixé :

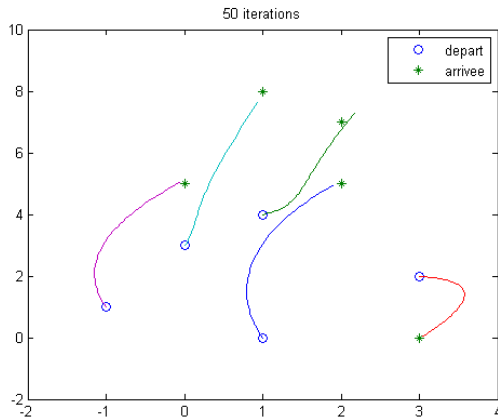
$$\begin{aligned} J(\alpha) &= \int_0^1 \|v_x(t)\|^2 dt + \mu \sum_{i=1}^N |y_i - x_i(1)|^2 \\ &= \int_0^1 \alpha^t S_V(x) \alpha dt + \mu \sum_{i=1}^N |y_i - x_i(1)|^2 \\ &= \int_0^1 C(x, \alpha) dt + g(x(1)) \end{aligned}$$

- Introduction d'un multiplicateur de Lagrange  $p(t)$
- Détour par le contrôle optimal et les équations hamiltoniennes :

$$\left\{ \begin{array}{l} \dot{x} = \frac{\partial H}{\partial p}(x, p, \alpha) \quad x(0) \text{ fixé} \\ \dot{p} = -\frac{\partial H}{\partial x}(x, p, \alpha) \quad p(1) = -Dg(x(1)) \\ \frac{\partial H}{\partial \alpha}(x, p, \alpha) = 0 \end{array} \right.$$

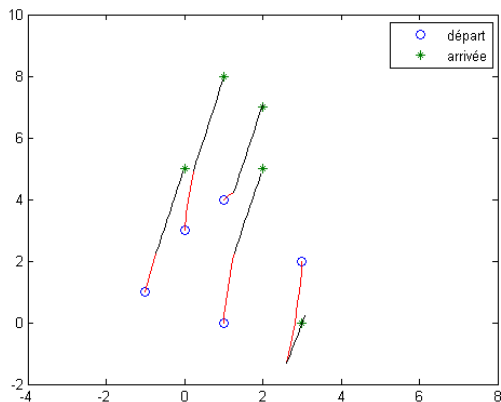




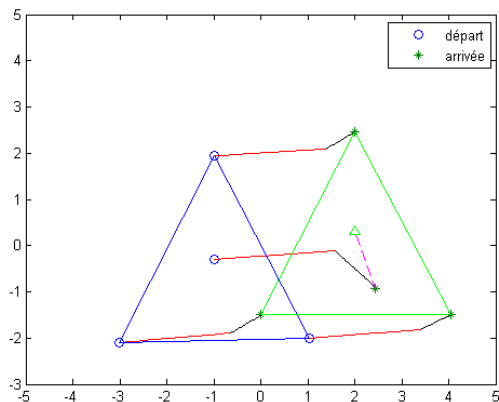


# Approche multi-échelles

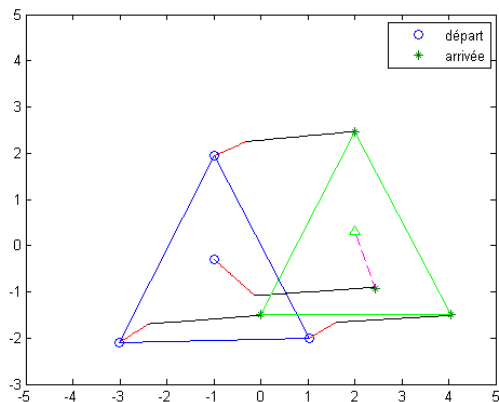
- On fait varier au cours du temps le paramètre d'échelle
- Les équations restent les mêmes, mais avec  $\lambda = \lambda(t)$
- Cela permet de séparer les composantes des différentes échelles de la transformation



(i)  $\lambda_1 = 0.1$  ;  $\lambda_2 = 10$  ;  $\mu = 20$



(j)  $\lambda_1 = 6$  ;  $\lambda_2 = 0.1$  ;  $\mu = 20$



(k)  $\lambda_1 = 0.1$  ;  $\lambda_2 = 6$  ;  $\mu = 100$

# Plan

- 1 Introduction
- 2 Méthodes linéaires
- 3 Méthode non linéaire
- 4 Moyenne de Karcher
- 5 Application : étude d'une maladie génétique chez les souris.
- 6 Conclusion

# Moyenne de Karcher

- $(x^k)_{1 \leq k \leq m}$  une population constituée de  $m$  formes
- Trouver un individu de référence
- Moyenne arithmétique :

$$\bar{x} = \frac{1}{m} \sum_{k=1}^m x^k$$

- Moyenne de Karcher :

$$\bar{x} = \operatorname{argmin}_y \sum_{k=1}^m d(y, x^k)^2$$

- Distance géodésique :

$$d(y, z) = \min_{\alpha} \left( \int_0^1 \alpha^t S_{V_{\lambda}}(x) \alpha dt \right) \quad x(0) = y \text{ et } x(1) = z$$

- Pseudo-distance géodésique :

$$\begin{aligned} d_1(y, z) &= \min_{\alpha} J(\alpha, y, z) \\ &= \min_{\alpha} \left( \int_0^1 \alpha^t S_{V_{\lambda}}(x) \alpha dt + \mu \sum_{i=1}^N |z_i - x_i(1)|^2 \right) \\ &\quad \text{avec } x(0) = y \end{aligned}$$

- Moyenne de Karcher :

$$J_0((\alpha^k)_k, \bar{x}, (x^k)_k) = \min_{x_0, (\alpha^k)_k} \sum_{k=1}^m J(\alpha^k, x_0, x^k)$$

$$= \min_{x_0, (\alpha^k)_k} \sum_{k=1}^m \left( \int_0^1 (\alpha^k)^t S_{V_\lambda}({}^k x) \alpha^k dt + \mu \sum_{i=1}^N |x_i^k - {}^k x_i(1)|^2 \right)$$

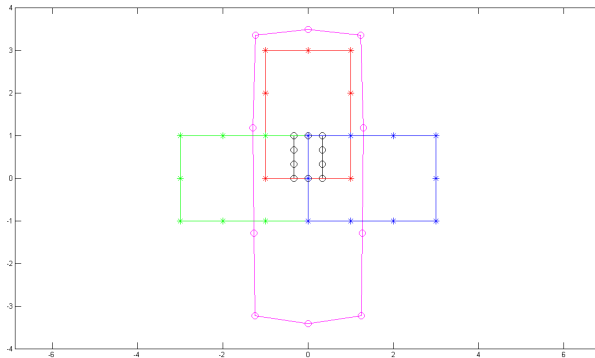
- Généralisation au cas multi-échelle :

$$J(\alpha, y, z) = \sum_{k=1}^m \int_{t_{k-1}}^{t_k} \alpha^t S_{V_{\lambda_k}}({}^k x) \alpha dt + \mu \sum_{i=1}^N |z_i - {}^k x_i(1)|^2$$

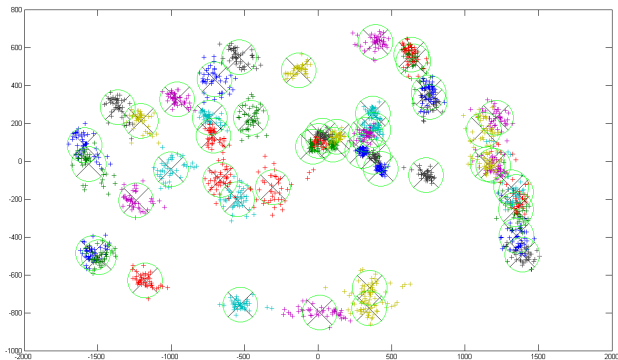
- Descente de gradient à la fois sur  $\alpha^k$  et sur  $x_0$
- Gradient sur les trajectoires :  $\alpha^k - p^k$
- Gradient sur la moyenne :  $-\sum_{k=1}^m p^k(0)$

Application : étude d'une maladie génétique chez les souris.

Conclusion



**FIGURE:** Exemple où la moyenne de Karcher diffère de la moyenne arithmétique pour les paramètres  $\lambda = 4$  et  $\mu = 100$



**FIGURE:** Ensemble des individus, moyenne de Karcher et moyenne arithmétique,  $\lambda = (46.9, 93.8, 187.5, 375, 750, 1500)$

# Plan

- 1 Introduction
- 2 Méthodes linéaires
- 3 Méthode non linéaire
- 4 Moyenne de Karcher
- 5 Application : étude d'une maladie génétique chez les souris.
  - Cadre de l'étude
  - Analyse statistique des résultats

## Cadre de l'étude

Données initiales :

- 44 souris (12 mutantes, 16 hétérozygotes, 16 saines)
- 49 landmarks par souris

But :

- Savoir à quelles échelles la mutation agit

Données obtenues après utilisation de nos programmes :

- Pour chacune des 6 échelles, 44 vecteurs  $X_i$  de  $\mathbb{R}^{196}$  (positions et moments des 49 landmarks).

# Analyse en Composantes Principales

But : obtenir le meilleur plan de projection pour visualiser des données de  $\mathbb{R}^N$ ,  $N \geq 3$ .

Méthode :

- Création de la matrice de Gram  $G$  des données obtenues :

$$G_{i,j} = \sum_{p=1}^N \sum_{q=1}^N \alpha_{i,p}^t K_V(x_i, x_j) \alpha_{j,q}$$

- Diagonalisation de  $G$  et récupération des 2 plus grandes valeurs propres ;
- Projection sur le plan engendré par les vecteurs propres correspondants.

Application : étude d'une maladie génétique chez les souris.

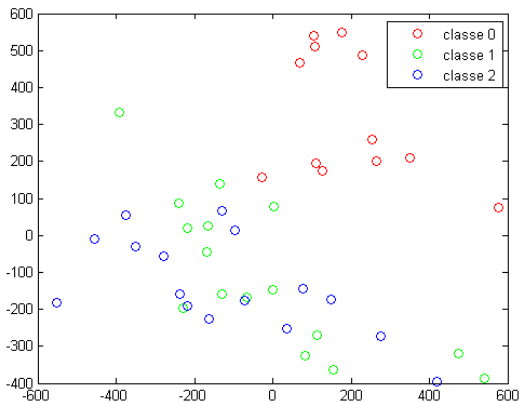
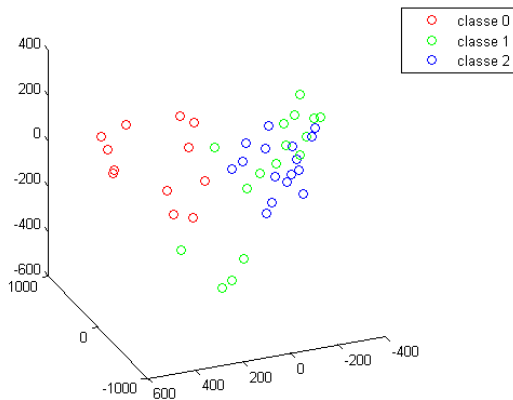


FIGURE: Projection sur 2 axes principaux

Application : étude d'une maladie génétique chez les souris.

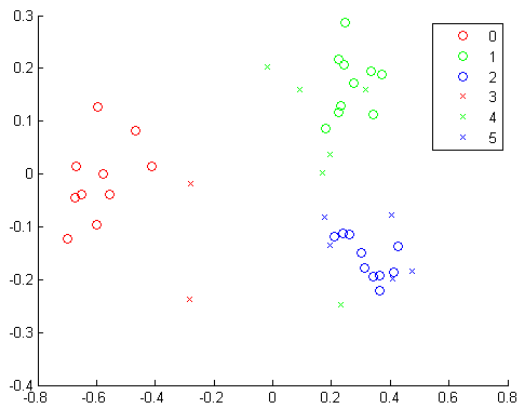
# ACP 3D



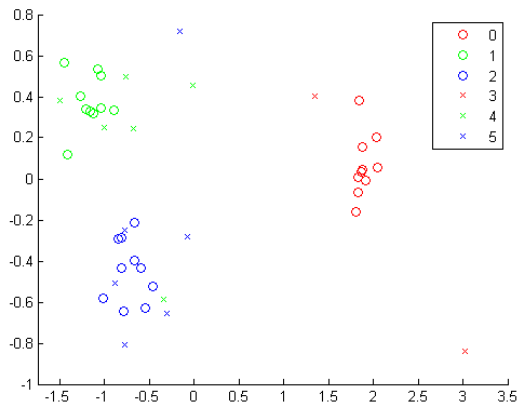
- $G$  matrice de Gram des données
- $\eta_i = G^{1/2} e_i$
- $j : X_i \rightarrow \eta_i$  est un plongement isométrique de  $\text{Vect}\{(X_i)\}$  dans  $\mathbb{R}^N$
- On peut alors appliquer la LDA

# Analyse de Fisher discriminante

- $S_W$  et  $S_B$  les matrices de covariance intra-classes et inter-classes
- On cherche :  $\max_w \frac{(S_B w, w)}{(S_W w, w)}$
- $S_B v = \lambda S_W v$
- Les vecteurs de projection correspondent aux plus grandes valeurs propres
- Si  $S_W$  n'est pas inversible, on lui ajoute  $\sigma^2 Id$  (régularisation)



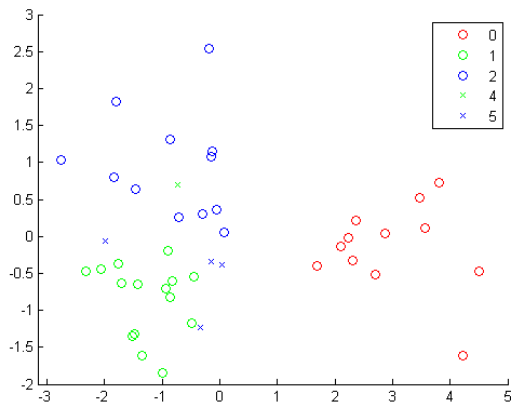
(a)  $\lambda = 0.25mm$  ;  $\mu = 10$



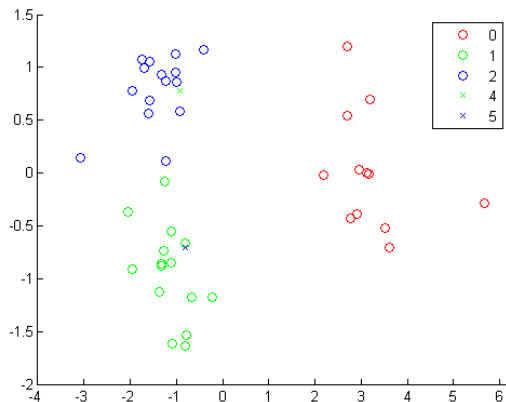
(b)  $\lambda = 8mm$  ;  $\mu = 10$

# Leave-one-out

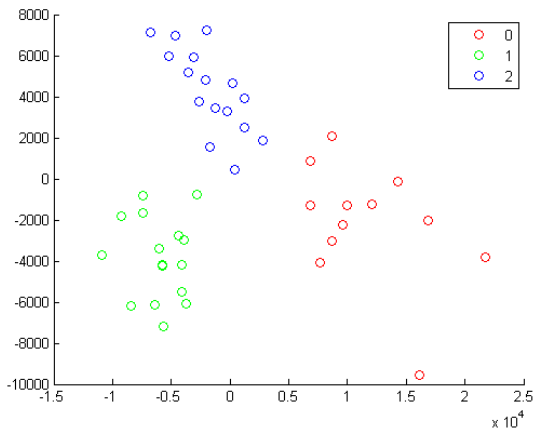
- On utilise toutes les données sauf une pour trouver le plan de projection
- On classe la dernière donnée
- On recommence en changeant la donnée de test



(c)  $\lambda = 4mm$  ;  $\mu = 10$



(d)  $\lambda = 0.5mm$  ;  $\mu = 10$



(e)  $\lambda = 2mm$  ;  $\mu = 10$

# Test de permutation

- Sert à vérifier la crédibilité d'un algorithme de classification
- On permute les étiquettes des individus
- On regarde le nombre d'erreurs de l'algorithme

Application : étude d'une maladie génétique chez les souris.

| Echelle           | Résultats réels | Meilleure permutation |
|-------------------|-----------------|-----------------------|
| $\lambda=0.25$ mm | 5               | 22                    |
| $\lambda=0.5$ mm  | 2               | 25                    |
| $\lambda=1$ mm    | 1               | 23                    |
| $\lambda=2$ mm    | 0               | 26                    |
| $\lambda=4$ mm    | 5               | 22                    |
| $\lambda=8$ mm    | 7               | 20                    |

# Plan

- 1 Introduction
- 2 Méthodes linéaires
- 3 Méthode non linéaire
- 4 Moyenne de Karcher
- 5 Application : étude d'une maladie génétique chez les souris.
- 6 Conclusion

# Conclusion

- Réduction du nombre de paramètres d'échelle.
- Choix d'un autre noyau vectoriel défini positif radial.
- Utiliser les programmes avec d'autres jeux de données.

# Conclusion

- Implémentation de programmes qui ont servi de briques.
- Test sur des exemples simples puis sur des données plus conséquentes.
- Analyse des résultats obtenus.