Séminaire MaMuX

Introduction aux outils de la géométrie de l'information pour la manipulation des flux audio

Arnaud Dessein, Arshia Cont, Gérard Assayag

10 octobre 2009

Plan

Cadre de la géométrie de l'information

- Géométrie de l'information :
  - Branche récente des mathématiques, en particulier de l'inférence statistique.
  - Etude des notions de probabilité et d'information par le biais de la géométrie différentielle.

Cadre de la géométrie de l'information

- Géométrie de l'information :
  - Branche récente des mathématiques, en particulier de l'inférence statistique.
  - Etude des notions de probabilité et d'information par le biais de la géométrie différentielle.
- Bref historique :
  - 1945 : Cramer et Rao, métrique d'information de Fisher.
  - Années 60 et 70 : Chentsov, métrique d'information de Fisher et connexions affines.
  - Années 80 : Amari et Nagaoka, connexions affines duales et divergences duales.

Cadre de la géométrie de l'information

- Géométrie de l'information :
  - Branche récente des mathématiques, en particulier de l'inférence statistique.
  - Etude des notions de probabilité et d'information par le biais de la géométrie différentielle.
- Bref historique :
  - 1945 : Cramer et Rao, métrique d'information de Fisher.
  - Années 60 et 70 : Chentsov, métrique d'information de Fisher et connexions affines.
  - Années 80 : Amari et Nagaoka, connexions affines duales et divergences duales.
- Livre de référence : Amari, S. & Nagaoka, H. (2000). *Methods of Information Geometry, volume 191 of Translations of Mathematical Monographs*. American Mathematical Society.

Concepts de base

- Point de départ : surface généralisée où chaque point est une distribution de probabilités.
  - Variété statistique : modèle statistique paramétrique $S = \{p_\xi = p(x;\xi) \colon \xi = [\xi^1, \dots, \xi^n] \in \Xi\}$.
  - Point de S : distribution de probabilités $p_\xi \colon x \mapsto p_\xi(x)$ pour $x \in \mathcal{X}$.
  - Coordonnées : paramètres ξ de p_ξ dans le modèle S.

Concepts de base

- Point de départ : surface généralisée où chaque point est une distribution de probabilités.
  - Variété statistique : modèle statistique paramétrique $S = \left\{p_\xi = p(x;\xi) \colon \xi = [\xi^1, \ldots, \xi^n] \in \Xi\right\}$.
  - Point de S : distribution de probabilités $p_\xi : x \mapsto p_\xi(x)$ pour $x \in \mathcal{X}$.
  - Coordonnées : paramètres ξ de p_ξ dans le modèle S.
- Exemple : S est la famille des distributions normales $p(x;\xi) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ avec $\xi = [\mu, \sigma]$.

Concepts de base

- Point de départ : surface généralisée où chaque point est une distribution de probabilités.
  - Variété statistique : modèle statistique paramétrique $S = \left\{p_\xi = p(x;\xi) : \xi = [\xi^1, \ldots, \xi^n] \in \Xi\right\}$.
  - Point de S : distribution de probabilités $p_\xi : x \mapsto p_\xi(x)$ pour $x \in \mathcal{X}$.
  - Coordonnées : paramètres ξ de p_ξ dans le modèle S.
- Exemple : S est la famille des distributions normales $p(x;\xi) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ avec $\xi = [\mu, \sigma]$.
- Géométrie intrinsèque de S ? Distance pertinente entre p_ξ et p_θ ?

Motivations musicales

- Structures de manipulation alternatives respectant les natures temporelle et probabiliste des flux audio.

Motivations musicales

- Structures de manipulation alternatives respectant les natures temporelle et probabiliste des flux audio.
- Propriétés géométriques intrinsèques permettant de définir une notion de similarité.

Motivations musicales

- Structures de manipulation alternatives respectant les natures temporelle et probabiliste des flux audio.
- Propriétés géométriques intrinsèques permettant de définir une notion de similarité.
- Applications audio :
  - Analyse des contenus audio : apprentissage automatique de structures, segmentation automatique, reconnaissance automatique de scènes sonores, etc.
  - Transformation des flux audio : restauration d'enregistrements, encodage et compression des données, nouvelles voies de transformation des sons dans un schéma d'analyse-synthèse, etc.

Motivations musicales

- Structures de manipulation alternatives respectant les natures temporelle et probabiliste des flux audio.
- Propriétés géométriques intrinsèques permettant de définir une notion de similarité.
- Applications audio :
  - Analyse des contenus audio : apprentissage automatique de structures, segmentation automatique, reconnaissance automatique de scènes sonores, etc.
  - Transformation des flux audio : restauration d'enregistrements, encodage et compression des données, nouvelles voies de transformation des sons dans un schéma d'analyse-synthèse, etc.
- Pour la musique :
  - Aide à l'analyse musicale.
  - Improvisation assistée par ordinateur.
  - Analyse, transformation, synthèse, recherche de sons.
  - Musique mixte et interactive aux temps de la composition et de la performance.

Plan

Variété topologique

Définitions.

- Une *variété topologique* de dimension n est un espace topologique séparé M localement homéomorphe à $\mathbb{R}^n$.
- Le couple (U, ϕ) est une *carte locale* de M.
- Une famille de cartes locales $\{(U_i, \phi_i)\}$ telle que $\bigcup_i U_i = M$ est un *atlas* de M.

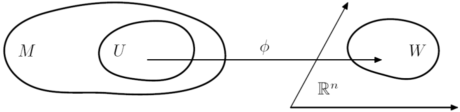

FIGURE: Variété topologique et carte locale.

Variété différentiable

Définition.

Une *variété différentiable* est une variété topologique M possédant un atlas $\{(U_i, \phi_i)\}$ tel que pour tous i, j avec $U_i \cap U_j \neq \emptyset$, l'application $\phi_j \circ \phi_i^{-1} \colon \phi_i(U_i \cap U_j) \to \phi_j(U_i \cap U_j)$ est différentiable.

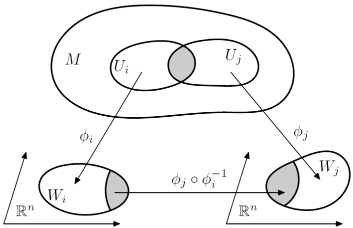

FIGURE: Variété différentiable.

Variété riemannienne

- Au point $p \in M$, on peut définir un espace vectoriel $T_p(M)$ de dimension n appelé *espace tangent* à M en p.

Variété riemannienne

- Au point $p \in M$, on peut définir un espace vectoriel $T_p(M)$ de dimension n appelé *espace tangent* à M en p.
- Sous certaines conditions, on peut munir tous les $T_p(M)$ de produits scalaires $\langle\ ,\ \rangle_p$ qui forment une *métrique riemannienne* de M.

Variété riemannienne

- Au point $p \in M$, on peut définir un espace vectoriel $T_p(M)$ de dimension n appelé *espace tangent* à M en p.
- Sous certaines conditions, on peut munir tous les $T_p(M)$ de produits scalaires $\langle \, , \rangle_p$ qui forment une *métrique riemannienne* de M.

Définition.

Une *variété riemannienne* est une variété différentiable M munie d'une métrique riemannienne.

Variété riemannienne

- Au point $p \in M$, on peut définir un espace vectoriel $T_p(M)$ de dimension n appelé *espace tangent* à M en p.
- Sous certaines conditions, on peut munir tous les $T_p(M)$ de produits scalaires $\langle\ ,\ \rangle_p$ qui forment une *métrique riemannienne* de M.

Définition.

Une *variété riemannienne* est une variété différentiable M munie d'une métrique riemannienne.

- Les structures engendrées par les espaces tangents et la métrique riemannienne sont locales et donnent accès à la notion de longueur d'une courbe.

Variété riemannienne

- Au point $p \in M$, on peut définir un espace vectoriel $T_p(M)$ de dimension n appelé *espace tangent* à M en p.
- Sous certaines conditions, on peut munir tous les $T_p(M)$ de produits scalaires $\langle \, , \rangle_p$ qui forment une *métrique riemannienne* de M.

Définition.

Une *variété riemannienne* est une variété différentiable M munie d'une métrique riemannienne.

- Les structures engendrées par les espaces tangents et la métrique riemannienne sont locales et donnent accès à la notion de longueur d'une courbe.
- On peut munir une variété riemannienne d'une *connexion affine* qui relie ces structures locales et donne accès aux notions de torsion, de courbure et de géodésique.

Plan

Modèles statistiques

- On considère un modèle statistique paramétrique de distributions de probabilités sur $\mathcal{X}$: $S = \left\{ p_\xi = p(x;\xi) \colon \xi = [\xi^1, \ldots, \xi^n] \in \Xi \right\}$.

Modèles statistiques

- On considère un modèle statistique paramétrique de distributions de probabilités sur $\mathcal{X}$: $S = \left\{p_\xi = p(x;\xi) \colon \xi = [\xi^1, \ldots, \xi^n] \in \Xi\right\}$.
- Hypothèses :
  - $p(x, \xi) > 0$ sur $\mathcal{X}$ pour tout $\xi \in \Xi$.
  - Ξ est un ouvert de $\mathbb{R}^n$.
  - Les p_ξ sont différentiables sur Ξ en tout point $x \in \mathcal{X}$.
  - On peut intervertir les ordres d'intégration et de différentiation.

Modèles statistiques

- On considère un modèle statistique paramétrique de distributions de probabilités sur $\mathcal{X}$: $S = \left\{ p_\xi = p(x;\xi) \colon \xi = [\xi^1, \dots, \xi^n] \in \Xi \right\}$.
- Hypothèses :
  - $p(x, \xi) > 0$ sur $\mathcal{X}$ pour tout $\xi \in \Xi$.
  - Ξ est un ouvert de $\mathbb{R}^n$.
  - Les p_ξ sont différentiables sur Ξ en tout point $x \in \mathcal{X}$.
  - On peut intervertir les ordres d'intégration et de différentiation.
- S est une variété différentiable de dimension n avec une carte globale (S, ϕ) où $\phi \colon p_\xi \mapsto \xi$. S est alors appelée *variété statistique*.

Modèles statistiques

- On considère un modèle statistique paramétrique de distributions de probabilités sur $\mathcal{X}$: $S = \left\{p_\xi = p(x;\xi) \colon \xi = [\xi^1, \dots, \xi^n] \in \Xi\right\}$.
- Hypothèses :
  - $p(x,\xi) > 0$ sur $\mathcal{X}$ pour tout $\xi \in \Xi$.
  - Ξ est un ouvert de $\mathbb{R}^n$.
  - Les p_ξ sont différentiables sur Ξ en tout point $x \in \mathcal{X}$.
  - On peut intervertir les ordres d'intégration et de différentiation.
- S est une variété différentiable de dimension n avec une carte globale (S, ϕ) où $\phi \colon p_\xi \mapsto \xi$. S est alors appelée *variété statistique*.
- Exemples : distribution normale, distribution normale multivariée, distribution de Poisson.

Métrique d'information de Fisher

Définition.

La *matrice d'information de Fisher* de S en ξ est la matrice semi-définie positive notée $G(\xi)$ telle que $g_{ij}(\xi) = E_\xi[\partial_i \ell_\xi \partial_j \ell_\xi] = \int \partial_i \ell(x;\xi) \partial_j \ell(x;\xi) p(x;\xi) \mathrm{d}x$ où $\partial_i = \frac{\partial}{\partial \xi_i}$ et $\ell_\xi(x) = \ell(x,\xi) = \log p(x;\xi)$.

Métrique d'information de Fisher

Définition.

La *matrice d'information de Fisher* de S en ξ est la matrice semi-définie positive notée $G(\xi)$ telle que $g_{ij}(\xi) = E_\xi[\partial_i \ell_\xi \partial_j \ell_\xi] = \int \partial_i \ell(x;\xi) \partial_j \ell(x;\xi) p(x;\xi) \mathrm{d}x$ où $\partial_i = \frac{\partial}{\partial \xi_i}$ et $\ell_\xi(x) = \ell(x,\xi) = \log p(x;\xi)$.

- Hypothèses :
  - $g_{ij}(\xi) < \infty$ pour tous i, j, ξ.
  - $g_{ij} : \Xi \to \mathbb{R}$ est différentiable pour tous i, j.
  - $G(\xi)$ est définie positive pour tout ξ.

Métrique d'information de Fisher

Définition.

La *matrice d'information de Fisher* de S en ξ est la matrice semi-définie positive notée $G(\xi)$ telle que $g_{ij}(\xi) = E_\xi[\partial_i \ell_\xi \partial_j \ell_\xi] = \int \partial_i \ell(x;\xi) \partial_j \ell(x;\xi) p(x;\xi) \mathrm{d}x$ où $\partial_i = \frac{\partial}{\partial \xi_i}$ et $\ell_\xi(x) = \ell(x,\xi) = \log p(x;\xi)$.

- Hypothèses :
  - $g_{ij}(\xi) < \infty$ pour tous i, j, ξ.
  - $g_{ij} : \Xi \to \mathbb{R}$ est différentiable pour tous i, j.
  - $G(\xi)$ est définie positive pour tout ξ.

Théorème (Cramer et Rao, Chentsov).

La matrice d'information de Fisher définit une métrique riemannienne g sur S. On appelle g la *métrique d'information de Fisher*. Sous certaines conditions, la métrique d'information de Fisher est l'unique métrique riemannienne sur S (à un facteur multiplicatif près).

Connexions affines alpha

Théorème (Chentsov).

Il existe une famille de connexions affines $\{\nabla^{(\alpha)}\}_{\alpha}$ sur (S, g). Cette famille est paramétrable par $\alpha \in \mathbb{R}$ et est unique (à un facteur multiplicatif près). Les $\nabla^{(\alpha)}$ sont appelées les *connexions affines* α.

Connexions affines alpha

Théorème (Chentsov).

Il existe une famille de connexions affines $\{\nabla^{(\alpha)}\}_\alpha$ sur (S, g). Cette famille est paramétrable par $\alpha \in \mathbb{R}$ et est unique (à un facteur multiplicatif près). Les $\nabla^{(\alpha)}$ sont appelées les *connexions affines* α.

- Il existe une notion de *dualité* entre $\nabla^{(\alpha)}$ et $\nabla^{(-\alpha)}$ (à un reparamétrage près) qui sont appelées *connexions affines duales* par rapport à g. La connexion affine duale d'une connexion affine ∇ est notée $\nabla^\star$.

Divergences

Définition.

Une *divergence* sur S est une fonction $D\colon S \times S \to \mathbb{R}$ telle que pour tous $p, q \in S$, $D(p \parallel q) \geqslant 0$ et $D(p \parallel q) = 0$ ssi $p = q$.

Divergences

Définition.

Une *divergence* sur S est une fonction $D \colon S \times S \to \mathbb{R}$ telle que pour tous $p, q \in S$, $D(p \parallel q) \geqslant 0$ et $D(p \parallel q) = 0$ ssi $p = q$.

Définition.

La *divergence duale* d'une divergence D est la divergence $D^\star$ définie par $D^\star(p \parallel q) = D(q \parallel p)$ pour tous p, q.

Divergences

Définition.

Une *divergence* sur S est une fonction $D\colon S \times S \to \mathbb{R}$ telle que pour tous $p, q \in S$, $D(p \parallel q) \geqslant 0$ et $D(p \parallel q) = 0$ ssi $p = q$.

Définition.

La *divergence duale* d'une divergence D est la divergence $D^\star$ définie par $D^\star(p \parallel q) = D(q \parallel p)$ pour tous p, q.

Théorème (Amari et Nagaoka).

Sous certaines conditions, une famille de connexions affines duales $\{(\nabla, \nabla^\star)\}$ sur (S, g) définit une unique famille de divergences duales $\{(D, D^\star)\}$. Réciproquement, une famille de divergences duales $\{(D, D^\star)\}$ sur S définit une unique métrique g et une unique famille de connexions affines duales $\{(\nabla, \nabla^\star)\}$.

Plan

Conclusion

- On peut définir une structure duale $(S, g, \nabla, \nabla^\star)$ de variété riemannienne sur une famille paramétrique de distributions de probabilités.

Conclusion

- On peut définir une structure duale $(S, g, \nabla, \nabla^\star)$ de variété riemannienne sur une famille paramétrique de distributions de probabilités.
- En représentant les flux audio dans une telle structure, on peut respecter leurs natures temporelle et probabiliste.

Conclusion

- On peut définir une structure duale $(S, g, \nabla, \nabla^\star)$ de variété riemannienne sur une famille paramétrique de distributions de probabilités.
- En représentant les flux audio dans une telle structure, on peut respecter leurs natures temporelle et probabiliste.
- Les propriétés géométriques intrinsèques sous-jacentes permettent de définir une notion de similarité par les divergences duales $D, D^\star$ associées aux connexions affines duales $\nabla, \nabla^\star$.

Conclusion

- On peut définir une structure duale $(S, g, \nabla, \nabla^\star)$ de variété riemannienne sur une famille paramétrique de distributions de probabilités.
- En représentant les flux audio dans une telle structure, on peut respecter leurs natures temporelle et probabiliste.
- Les propriétés géométriques intrinsèques sous-jacentes permettent de définir une notion de similarité par les divergences duales $D, D^\star$ associées aux connexions affines duales $\nabla, \nabla^\star$.
- De nombreuses applications à l'audio et à la musique en particulier sont envisageables.