Journées MASCOT NUM

22-23 mars 07

Introduction aux RKHS

L. Carraro, X. Bay, O. Roustant

Ecole des mines de Saint-Etienne

(nom@emse.fr)

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

PLAN

- Partie 1 (22 mars) :
  - Motivation : régularisation
  - Noyaux de Bergmann
  - RKHS, définition et propriétés
  - Exemples
  - Retour sur la régularisation

PLAN

- Partie 2 (23 mars) :
  - RKHS et régularisation, exemples.
  - Le noyau reproduisant et ses interprétations.
  - Discrépance et RKHS
  - Propagation d'incertitudes et RKHS
  - Semi-RKHS

PLAN

- Partie 1 (22 mars) :
  - Motivation : régularisation
  - Noyaux de Bergmann
  - RKHS, définition et propriétés
  - Exemples
  - Retour sur la régularisation

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

Régularisation

Estimation non paramétrique :

données = couples (x_i, y_i) pour $1 \leq i \leq n$

modèle :

$$y_i = f(x_i) + \varepsilon_i, \quad où \quad \varepsilon_i \ v.a.i.i.d. \ N(0, \sigma^2)$$

estimation de f ?

$$-2 \log L(f) \propto (1/\sigma^2) \sum_{i=1}^{n} \left(y_i - f(x_i) \right)^2$$

Problème mal posé :

- Pénaliser
- Informations a priori

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

Régularisation

Vraisemblance pénalisée :

critère à minimiser :

$$C(f) = \boldsymbol{L}(\boldsymbol{y}, f(\boldsymbol{x})) + \lambda \ J(f)$$

- J est une fonctionnelle mesurant la régularité de f
- λ est un paramètre de lissage
- $$\boldsymbol{L}(\boldsymbol{y}, f(\boldsymbol{x})) = \sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2$$

Plus généralement,

- $$\boldsymbol{L}(\boldsymbol{y}, f(\boldsymbol{x})) = -2 \ \log L(f), \ \ ou \ \ \boxed{\boldsymbol{L}(\boldsymbol{y}, f(\boldsymbol{x})) = \sum_{i=1}^{n} D(y_i, f(x_i))}$$

D mesure une distance entre y_i et $f(x_i)$

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

Cas limites

- $\lambda \to +\infty$ $\quad \underset{J(f)=0}{Min} \sum_{i=1}^{n} D(y_i, f(x_i))$

Souvent, {f, J(f) = 0} est un e.v. de dim finie

→ régression linéaire

- $\lambda \to 0$ $\quad \underset{f(x_i)=y_i}{Min} J(f)$

→ interpolation

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

PLAN

- Partie 1 (22 mars) :
  - Motivation : régularisation
  - Noyaux de Bergmann
  - RKHS, définition et propriétés
  - Exemples
  - Retour sur la régularisation

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

Noyaux de Bergmann

Un espace de Hilbert et un noyau

U ouvert de $\mathbb{C}$

H = {f holomorphe sur U, $\iint_U |f(u)|^2 du < +\infty$}

H est un Hilbert avec :

$$\langle f, g \rangle = \iint_U f(u)\overline{g(u)}du$$

Alors : f → f(v) est continue, donc (Riesz) :

$$f(v) = \delta_v(f) = \langle f, K_v \rangle$$

Noyau de Bergmann

Notation :

$$K(u,v) := \overline{K_v(u)}$$

K est le noyau de Bergmann de l'ouvert U.

$$f(u) = \iint_U f(v)\, K(u,v)\, dv$$

C'est un résultat de représentation intégrale.

Exemple pour U = D(0,1[:

$$K(u,v) = \frac{\pi}{(1-\bar{u}\,v)^2}$$

Ecole Nationale Supérieure des Mines
SAINT-ETIENNE

PLAN

- Partie 1 (22 mars) :
  - Motivation : régularisation
  - Noyaux de Bergmann
  - RKHS, définition et propriétés
  - Exemples
  - Retour sur la régularisation

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

RKHS - définition

- Espace de Hilbert à noyau reproduisant

  (Reproducing Kernel Hilbert Space)

X est un ensemble d'indices :

$$X = \{1, \dots, n\}, \quad X = \mathbb{R}, \quad X = \mathbb{R}^d, \ \dots$$

$E = \mathbb{R}^X$ est muni de la topologie de la convergence simple.

H ⊂ E est un RKHS si :

- H est un espace de Hilbert, de produit scalaire $\langle f, g \rangle$
- L'injection H → E est continue

RKHS et noyau

Pour tout x de X, il existe K_x tel que :

$$f(x) = \delta_x(f) = \langle f, K_x \rangle$$

Notation :

$$K(x,y) := K_x(y)$$

Propriétés : K est un noyau

reproduisant : $K(x,y) = \langle K_x, K_y \rangle$

symétrique défini positif :

$$\forall a = (a_x)_x \in \mathbb{R}^{(X)}, \sum_x \sum_y a_x a_y K(x,y) \geq 0$$

$$et = 0 \quad ssi \quad \sum_x a_x K_x \equiv 0$$

RKHS et noyau

Réciproquement, si K est un noyau défini positif sur X, c'est le noyau reproduisant d'un RKHS

Idée de preuve.

On pose $K_x(y) = K(x,y)$ et on définit sur l'e.v. F engendré par les K_x un produit scalaire :

$$\langle \sum_x a_x K_x , \sum_y b_y K_y \rangle = \sum_x \sum_y a_x b_y K(x,y)$$

F est un espace préhilbertien sur lequel δ_x est continue.

On complète F en H, qui reste un sous-espace de E, et l'injection est continue.

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

RKHS et noyau

L'ensemble **K** des noyaux symétriques définis positifs est un cône convexe saillant :

- $K_1, K_2 \in \mathbf{K}, \lambda > 0 \Rightarrow \quad \lambda K_1 + K_2 \in \mathbf{K}$
- $\mathbf{K} \cap (-\mathbf{K}) = \{0\}$

¿

$K_{1,} K_2 \in K$

L'ensemble des RKHS inclus dans E est également un cône convexe saillant ; la bijection précédente conserve ces structures.

Voir Schwartz (1964).

PLAN

- Partie 1 (22 mars) :
  - Motivation : régularisation
  - Noyaux de Bergmann
  - RKHS, définition et propriétés
  - Exemples
  - Retour sur la régularisation

Exemple 1 : X={1,...,n}

Soit K noyau défini positif sur X×X. On note **K** la matrice de terme général K(x,y).

K défini un produit scalaire sur H selon :

$$\text{Si } u = \sum_x a_x K_x = \boldsymbol{K}\,\boldsymbol{a} \text{ et } v = \sum_x b_x K_x = \boldsymbol{K}\,\boldsymbol{b}$$

$$\langle u, v\rangle = \sum_x \sum_y a_x b_y K(x,y) = \boldsymbol{a}'\,\boldsymbol{K}\,\boldsymbol{b}$$

$$\text{Donc } \langle u, v\rangle = \boldsymbol{u}'\,\boldsymbol{K}^{-1}\,\boldsymbol{v}$$

Remarques :

$$\langle u, K_x\rangle = u(x)$$

$$\text{Si } \lambda = \sum_x a_x \delta_x \text{ et } \mu = \sum_x b_x \delta_x \text{ sont dans H'}, \langle \lambda, \mu\rangle = \boldsymbol{a}'\,\boldsymbol{K}\,\boldsymbol{b}$$

$$u \in H \rightarrow \langle u, .\rangle \in H' \text{ est une isométrie}$$

Exemple 1 : X={1,…,n}

Diagonalisation de la matrice **K** :

$\boldsymbol{K} = \boldsymbol{P}'\,\boldsymbol{\Lambda}\,\boldsymbol{P}$ avec $\boldsymbol{P}$ orthogonale et $\boldsymbol{\Lambda}$ diagonale

Donc $K(x,y) = \sum_{i=1}^{n} \lambda_i \varphi_i(x) \varphi_i(y)$

Et dans cette nouvelle base :

Si $u = \sum_{i=1}^{n} \alpha_i\,\varphi_i$ et $v = \sum_{i=1}^{n} \beta_i\,\varphi_i$

Et $\langle u,v \rangle = \sum_{i=1}^{n} \frac{\alpha_i \beta_i}{\lambda_i}$

Remarque :

Extension des ces résultats en dimension infinie.

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

Exemple 2 - X=[0,1]

$$H = H_0^1 = \{ f : [0,1] \to \mathbb{R}, f \in L^2, f' \in L^2, f(0) = 0 \}$$

$$\langle f, g \rangle = \int_0^1 f'(x) g'(x) dx$$

Sur H, δ_x est continue et :

$$f(x) = \int_0^x f'(y) dy = \int_0^1 f'(y) 1_{[0,x]}(y) dy$$

Donc si $K_x(0) = 0$ et $K_x' = 1_{[0,x]}$ i.e. $K(x,y) = min(x,y)$

K est le noyau reproduisant de H.

Exemple 2 - X=[0,1]

$$H = H^2_{00} = \{ f : [0,1] \to \mathbb{R}, f \in L^2, f' \in L^2, f'' \in L^2, f(0)=0, f'(0)=0 \}$$

$$\langle f, g \rangle = \int_0^1 f''(x) g''(x)\, dx$$

Sur H, δ_x est continue et :

$$f(x) = \int_0^x (x-y) f''(y)\, dy = \int_0^1 f''(y)(x-y)_+\, dy$$

D'où $K(x,y) = \dfrac{(x-y)_+^3}{6} + P_1(y) = -\dfrac{y^3}{6} + x\dfrac{y^2}{2}$, si $y \le x$

NB : K est symétrique

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

PLAN

- Partie 1 (22 mars) :
  - Motivation : régularisation
  - Noyaux de Bergmann
  - RKHS, définition et propriétés
  - Exemples
  - Retour sur la régularisation

RKHS et régularisation

- Théorème (Kimerldorf et Wahba 1970 ?)

Soit H un RKHS de noyau K. Le minimum du critère C, défini par :

$$C(f) = \boldsymbol{L}(\boldsymbol{y}, f(\boldsymbol{x})) = \sum_{i=1}^{n} D(y_i, f(x_i)) + \lambda \|f\|_H^2$$

est obtenu pour une fonction f de la forme :

$$f = \sum_{i=1}^{n} \beta_i \, K_{x_i}$$

Preuve

Début de preuve dans le cas d'un espace de Banach H, strictement convexe, où les évaluations sont continues.

On montre que pour toute fonction f de H, il existe g dans un « espace » de dimension au plus n telle que :

$$C(g) \leq C(f)$$

$$\text{Soit } F = \bigcap_i Ker(\delta_{x_i})$$

Soit P_F la projection sur F :

$$\|f - P_F(f)\| \leq \|f\| \text{ et } (f - P_F(f))(x_i) = f(x_i)$$

Donc $C(f - P_F(f)) \leq C(f)$

Preuve (retour à H Hilbert)

Soit G l'e.v. engendré par les fonctions $K(x_i,.)$

On note que : $G = F^{\perp}$

Par suite,

$$f - P_F(f) = P_G(f)$$

C'est tout !!

Commentaires

- La solution du pb variationnel vit dans un espace de dimension finie

Notations :

$$\Gamma = \left[K(x_i, x_j)\right]_{i,j} \; et \; \beta = [\beta_1 \ldots \beta_n]'$$

- Pb de minimisation :

$$\underset{\beta}{Min} \sum_{i=1}^{n} D(y_i, (\Gamma\beta)_i) + \lambda \beta' \Gamma \beta$$

Si D(y,z) = (y-z)²,

$$\underset{\beta}{Min} (y - \Gamma\beta)'(y - \Gamma\beta) + \lambda \beta' \Gamma \beta$$

Ecole Nationale Supérieure des Mines SAINT-ETIENNE

Commentaires

- Estimation de λ

  Validation croisée et variantes (GCV, GACV…)

- Extension à d'autres observations.

  Exemple : observation de moyennes sur des régions.

  Si L est une forme linéaire continue sur H :

$$L(f) = \langle f, \eta \rangle, \text{ où } \eta(x) = L(K_x)$$

- Extension à une semi-norme

$$C(f) = \sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2 + \lambda \int_0^1 f''(x)^2 \, dx$$