Noyaux définis positifs

Cours Master 2004/05

Jean-Philippe Vert

Jean-Philippe.Vert@mines.org

Plan

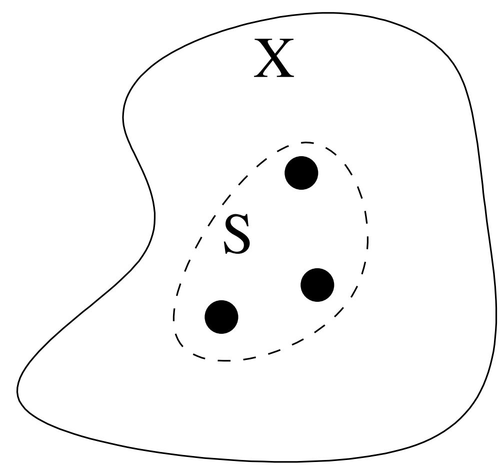
- Représentation des données par similarité
- Noyau défini positif
- Théorème de Mercer
- Espace de Hilbert à noyau reproduisant (rkhs)
- Noyau de Mercer et rkhs
- Rkhs et fonctions de Green
- Noyaux et régularisation par transformée de Fourier

Représentation des données par similarité

Les données

Soit $\mathcal{X}$ un ensemble. Un *objet* est un point $\mathbf{x} \in \mathcal{X}$. Un ensemble de données à analyser sera souvent un ensemble fini de N objets: $\mathcal{S} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$. Par exemple:

- $\mathcal{X}$ est l'ensemble de toutes les séquences finies dans l'alphabet $\{A, C, G, T\}$
- Un objet $\mathbf{x} \in \mathcal{X}$ est une séquence d'ADN qui code pour un gène
- $\mathcal{S}$ est l'ensemble est 30.000 gènes humains



Représentation classique des données

Soit un *algorithme* A capable de traîter des données d'un espace $\mathcal{F}$, par exemple $A : \mathcal{F}^N \mapsto \mathbb{R}$. Pour traîter $\mathcal{S}$, il faut soit $\mathcal{X} = \mathcal{F}$, soit définir une application $\Phi : \mathcal{X} \mapsto \mathcal{F}$ et travailler sur l'ensemble:

$$\Phi(\mathcal{S}) = \{\Phi(\mathbf{x}_1), \ldots, \Phi(\mathbf{x}_N)\} \in \mathcal{F}^N .$$

Par exemple:

- $\mathcal{F} =$ l'ensemble des séquences finies, pour un algorithme de compression
- $\mathcal{F} = \mathbb{R}^d$, pour un algorithme qui calcule le barycentre

Représentation par comparaison

Plutôt que de représenter chaque objet $\mathbf{x} \in \mathcal{X}$ individuellement par $\Phi(\mathbf{x}) \in \mathcal{F}$, et donc $\mathcal{S} \in \mathcal{X}^N$ par $\Phi(\mathcal{S}) \in \mathcal{F}^N$, soit une *fonction de similarité*:

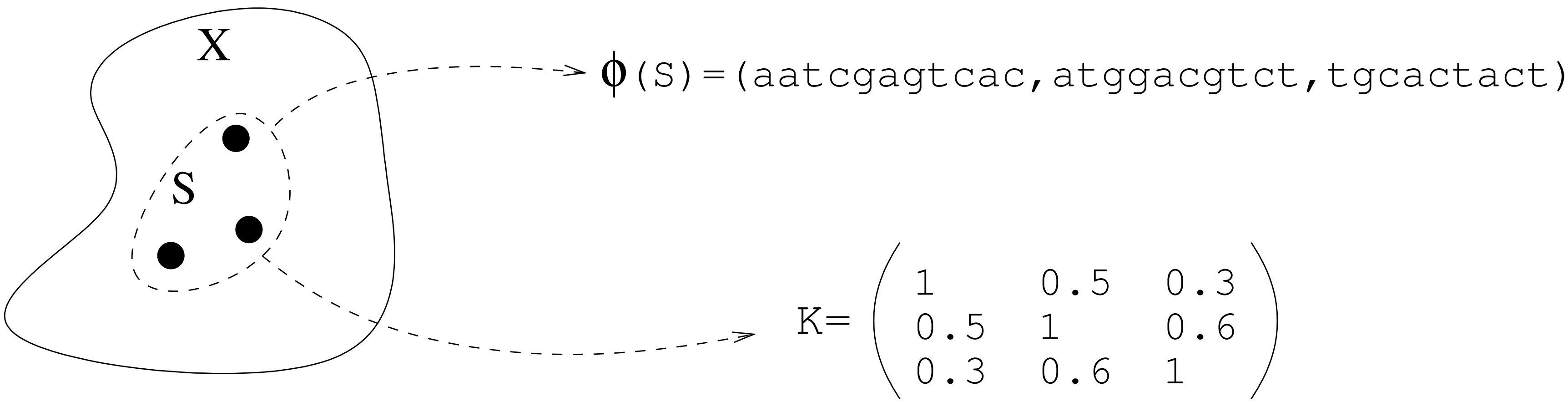
$$K : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}.$$

On peut alors représenter $\mathcal{S}$ par la *matrice de similarité* $N \times N$:

$$[K]_{ij} := K(\mathbf{x}_i, \mathbf{x}_j)$$

On utilisera alors des algorithmes capables de traîter des matrices carrées.

Les deux représentations



Remarques

- La représentation par similarité est *toujours une matrice carrée réelle*, quels que soient les objets (séquences, vecteurs, images, ...). Le même algorithme pourra traîter toutes ces données.
- Il y aura une complète *modularité* entre le choix de la fonction de similarité, d'une part, et le choix de l'algorithme qui sera appliqué à la matrice de similarité, d'autre part.
- La *taille* de la matrice de similarité est toujours $N \times N$, quelles que soient la nature et la complexité des objets
- Il est parfois *plus simple* de comparer des objets complexes que de les transformer en une forme imposée par un algorithme.

Noyaux définis positifs

Noyaux définis positifs (n.d.p.)

Nous nous restreindrons dans ce cours à une classe particulière de fonctions de similarité:

Définition 1 *Un noyau défini positif (n.d.p.) sur l'ensemble $\mathcal{X}$ est une fonction $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ symétrique:*

$$\forall \left(\mathbf{x}, \mathbf{x}'\right) \in \mathcal{X}^2, \quad K\left(\mathbf{x}, \mathbf{x}'\right) = K\left(\mathbf{x}', \mathbf{x}\right),$$

et qui satisfait, pour tout $N \in \mathbb{N}$, $\left(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\right) \in \mathcal{X}^N$ et $\left(a_1, a_2, \ldots, a_N\right) \in \mathbb{R}^N$:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j K\left(\mathbf{x}_i, \mathbf{x}_j\right) \geq 0.$$

N.d.p. (suite)

- Définition équivalente: pour tout ensemble d'objets $\mathcal{S} \in \mathcal{X}^N$, *la matrice de similarité est symétrique semi-définie positive*.
- En nous restreignant aux n.d.p., nous allons pouvoir utiliser une large famille d'algorithmes, les *méthodes à noyau*, qui travaillent sur des matrices symétriques semi-définies positives.

Le plus simple n.d.p.

Soit $\mathcal{X} = \mathbb{R}^d$ et la fonction $K : \mathcal{X}^2 \mapsto \mathbb{R}$ définie par:

$$\forall \left(\mathbf{x}, \mathbf{x}'\right) \in \mathcal{X}^2, \quad K\left(\mathbf{x}, \mathbf{x}'\right) = \left\langle \mathbf{x}, \mathbf{x}' \right\rangle$$

On vérifie que K est un n.d.p., appelé le *noyau linéaire*:

- $\langle \mathbf{x}, \mathbf{x}' \rangle = \langle \mathbf{x}', \mathbf{x} \rangle$,
- $\sum_{i=1}^N \sum_{j=1}^N a_i a_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle = \| \sum_{i=1}^N a_i \mathbf{x}_i \|^2 \geq 0$

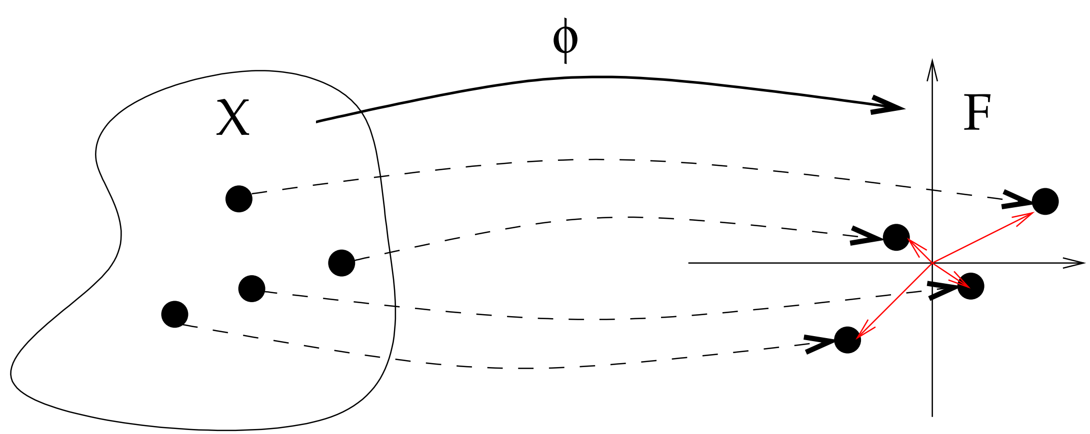
Un n.p.d. plus ambitieux

Soit $\mathcal{X}$ quelconque, et $\Phi : \mathcal{X} \mapsto \mathbb{R}^d$. Alors la fonction $K : \mathcal{X}^2 \mapsto \mathbb{R}$ définie par:

$$\forall \left(\mathbf{x}, \mathbf{x}'\right) \in \mathcal{X}^2, \quad K\left(\mathbf{x}, \mathbf{x}'\right) = \left\langle \Phi\left(\mathbf{x}\right), \Phi\left(\mathbf{x}'\right) \right\rangle$$

est un n.d.p, car:

- $\left\langle \Phi\left(\mathbf{x}\right), \Phi\left(\mathbf{x}'\right) \right\rangle = \left\langle \Phi\left(\mathbf{x}'\right), \Phi\left(\mathbf{x}\right) \right\rangle$,
- $\sum_{i=1}^N \sum_{j=1}^N a_i a_j \left\langle \Phi\left(\mathbf{x}_i\right), \Phi\left(\mathbf{x}_j\right) \right\rangle = \parallel \sum_{i=1}^N a_i \Phi\left(\mathbf{x}_i\right) \parallel^2 \geq 0$



Réciproquement...

Nous allons démontrer la réciproque suivante:

Théorème 2 *Si K est un n.p.d. sur un espace $\mathcal{X}$ quelconque, alors* il existe un espace de Hilbert $\mathcal{H}$ *muni du produit scalaire $\langle .,. \rangle_{\mathcal{H}}$ et une application*

$$\Phi : \mathcal{X} \mapsto \mathcal{H},$$

tels que:

$$\forall \left(\mathbf{x}, \mathbf{x}'\right) \in \mathcal{X}^2, \quad K\left(\mathbf{x}, \mathbf{x}'\right) = \left\langle \Phi\left(\mathbf{x}\right), \Phi\left(\mathbf{x}'\right) \right\rangle_{\mathcal{H}}.$$

Au cas où...

Définition 3

- *Un produit scalaire sur un $\mathbb{R}$-espace vectoriel $\mathcal{H}$ est une application $(f, g) \mapsto \langle f, g \rangle_{\mathcal{H}}$ de $\mathcal{H}^2$ dans $\mathbb{R}$ qui est bilinéaire, symétrique et telle que $\langle f, f \rangle > 0$ pour tout $f \in \mathcal{H} \backslash \{0\}$.*
- *Un espace vectoriel muni d'un produit scalaire est appelé pré-hilbertien. Il est muni d'une norme associée au produit scalaire par $\| f \|_{\mathcal{H}} = \langle f, f \rangle_{\mathcal{H}}^{\frac{1}{2}}$.*
- *Un espace de Hilbert est un espace vectoriel muni d'un produit scalaire et complet pour la norme associée.*

Preuves du Théorème 2

- Depuis l'origine de l'algèbre si $\mathcal{X}$ est fini
- Mercer (1909) pour $\mathcal{X} = [a, b] \subset \mathbb{R}$ et K continue
- Kolmogorov (1941) pour $\mathcal{X}$ dénombrable
- Aronszajn (1944, 1950) pour le cas général

Preuve: cas $\mathcal{X}$ fini

- On suppose $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ fini
- Si K est un n.d.p. alors la matrice $N \times N$ de similarité est *symétrique définie positive* donc *diagonalisable dans une base orthonormée* $(u_1, u_2, \ldots, u_N)$ avec des valeurs propres $0 \leq \lambda_1 \leq \ldots \leq \lambda_N$, ce qui s'écrit:

$$K\left(\mathbf{x}_i, \mathbf{x}_j\right) = \sum_{k=1}^{N} \lambda_k u_{i,k} u_{j,k} = \left\langle \Phi\left(\mathbf{x}_i\right), \Phi\left(\mathbf{x}_j\right)\right\rangle_{\mathbb{R}^N},$$

avec

$$\Phi\left(\mathbf{x}_i\right) = \begin{pmatrix} \sqrt{\lambda_1} u_{i,1} \\ \vdots \\ \sqrt{\lambda_N} u_{i,N} \end{pmatrix}$$

Théorème de Mercer

Noyau de Mercer

Nous allons prouver le Théorème dans le cas suivant:

- $\mathcal{X}$ est un *espace métrique compact* (typiquement, un fermé borné dans $\mathbb{R}^d$)
- K un n.d.p. *continu* sur $\mathcal{X} \times \mathcal{X}$ (par rapport à la tribu Borélienne)

Un tel noyau est appelé *noyau de Mercer*. La preuve repose sur la construction d'un opérateur linéaire compact et sa diagonalisation.

Rappel (au cas ou...)

Définition 4 *Soit $\mathcal{H}$ un espace de Hilbert*

- *Un opérateur linéaire est une application linéaire continue de $\mathcal{H}$ dans lui-même.*
- *On dit qu'un opérateur linéaire L est compact si pour toute suite bornée $\{f_n\}_{n=1}^{\infty}$, la séquence $\{Lf_n\}_{n=1}^{\infty}$ possède une sous-séquence convergente*
- *L est dit auto-adjoint si pour tout $f, g \in \mathcal{H}$:*

$$\langle f, Lg \rangle = \langle Lf, g \rangle .$$

- *L est dit positif ssi il est auto-adjoint et pour tout $f \in \mathcal{H}$:*

$$\langle f, Lf \rangle \geq 0$$

Un lemme important

- Soit ν un mesure de Borel sur $\mathcal{X}$, $L_2^\nu(\mathcal{X})$ l'espace de Hilbert des fonctions de carré intégrable
- Pour toute fonction $K : \mathcal{X}^2 \mapsto \mathbb{R}$, soit la tranformation L_K définie par:

$$\forall f \in L_2^\nu(\mathcal{X}), (L_K f)(\mathbf{x}) = \int K(\mathbf{x}, \mathbf{t}) f(\mathbf{t}) d\nu(\mathbf{t}).$$

Proposition 5 *Si K est un noyau de Mercer, alors L_K est un opérateur linéaire borné compact de $L_2^\nu(\mathcal{X})$, auto-adjoint et positif.*

Preuve

L_K est une application de $L_2^\nu(\mathcal{X})$ ans $L_2^\nu(\mathcal{X})$:
Pour tout $f \in L_2^\nu(\mathcal{X})$ et $(\mathbf{x}_1, \mathbf{x}_1) \in \mathcal{X}^2$:

$$
\begin{aligned}
\mid L_K f(\mathbf{x}_1) - L_K f(\mathbf{x}_2) \mid &= \left| \int (K(\mathbf{x}_1, \mathbf{t}) - K(\mathbf{x}_2, \mathbf{t})) f(\mathbf{t}) d\nu(\mathbf{t}) \right| \\
&\leq \parallel K(\mathbf{x}_1, \cdot) - K(\mathbf{x}_2, \cdot) \parallel \parallel f \parallel \text{ (Cauchy-Sc} \\
&\leq \sqrt{\nu(\mathcal{X})} \max_{\mathbf{t} \in \mathcal{X}} \mid K(\mathbf{x}_1, \mathbf{t}) - K(\mathbf{x}_2, \mathbf{t}) \mid \parallel f \parallel.
\end{aligned}
$$

K étant continu et $\mathcal{X}$ compact, K est uniformément continue,

donc *$L_K f$ est continue*. En particulier, $L_K f \in L_2^\nu(\mathcal{X})$. $\square$

Preuve

L_K est linéaire et continue
La linéarité est triviale.
D'autre part, on a $\forall f \in L_2^\nu(\mathcal{X})$ et $\forall \mathbf{x} \in \mathcal{X}$:

$$\begin{aligned} |(L_K f)(\mathbf{x})| &= \left| \int K(\mathbf{x}, \mathbf{t}) f(\mathbf{t}) d\nu(\mathbf{t}) \right| \\ &\leq \sqrt{\nu(\mathcal{X})} C_{K,x} \| f \|. \end{aligned}$$

avec $C_{K,x} = \max_{\mathbf{t} \in \mathcal{X}} |K(\mathbf{x}, \mathbf{t})|$. Donc:

$$\| L_K f \| = \left(\int L_K f(\mathbf{t})^2 d\nu(\mathbf{t}) \right)^{\frac{1}{2}} \leq \nu(\mathcal{X}) C_{K,x} \| f \|. \quad \square$$

Rappel: Théorème d'Ascoli

Soit $C(\mathcal{X})$ l'ensemble des fonctions continues sur $\mathcal{X}$, muni de la norme infinie $\| f \|_\infty = max_{\mathbf{x}\in\mathcal{X}} \mid f(\mathbf{x}) \mid$.
Un ensemble de fonctions $G \subset C(\mathcal{X})$ est dit *equicontinu* ssi:

$$\forall \epsilon > 0, \exists \delta > 0, \forall (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^2,$$
$$\| \mathbf{x} - \mathbf{y} \| < \delta \implies \forall g \in G, \mid g(\mathbf{x}) - g(\mathbf{y}) \mid < \epsilon.$$

Théorème 6 *(Ascoli) Une partie $H \subset C(\mathcal{X})$ est* relativement compacte *(i.e., son adhérence est compacte) ssi elle est* uniformément bornée *et* équicontinue*.*

Preuve

L_K est compact:
Soit $(f_n)_{n\geq 0}$ une séquence bornée de $L_2^{\nu}(\mathcal{X})$ ($\| f_n \| < M$ pour tout n).
La suite $(f_n)_{n\geq 0}$ est une suite de fonctions continues, uniformément bornée car:

$$\| L_K f \|_\infty \leq \sqrt{\nu(\mathcal{X})} C_K \| f \| \leq \sqrt{\nu(\mathcal{X})} C_K M.$$

Elle est d'autre part equicontinue, car

$$| L_K f_n(\mathbf{x}_1) - L_K f_n(\mathbf{x}_2) | \leq \sqrt{\nu(\mathcal{X})} \max_{\mathbf{t}\in\mathcal{X}} | K(\mathbf{x}_1, \mathbf{t}) - K(\mathbf{x}_2, \mathbf{t}) | M.$$

Par le Théorème d'Ascoli, on peut extraire une sous-suite uniformément convergente dans $C(\mathcal{X})$, et donc dans $L_2^{\nu}(\mathcal{X})$. $\square$

Preuve

L_K *est auto-adjoint:* K étant symétrique, alors on a pour tout $f, g \in \mathcal{H}$:

$$\begin{aligned}\langle f, Lg \rangle &= \int f(\mathbf{x})(Lg)(\mathbf{x})\,\nu(d\mathbf{x}) \\ &= \int\int f(\mathbf{x})\,g(\mathbf{t})\,K(\mathbf{x},\mathbf{t})\,\nu(d\mathbf{x})\,\nu(d\mathbf{t}) \quad \text{(Fubini)} \\ &= \langle Lf, g \rangle\,.\end{aligned}$$

Preuve

L_K *est positif:* On approxime l'intégrale par des sommes finies:

$$\begin{aligned}
\langle f, Lf \rangle &= \int\int f(\mathbf{x}) f(\mathbf{t}) K(\mathbf{x}, \mathbf{t}) \nu(d\mathbf{x}) \nu(d\mathbf{t}) \\
&= \lim_{k\to\infty} \frac{\nu(\mathcal{X})}{k^2} \sum_{i,j=1}^{k} K(\mathbf{x}_i, \mathbf{x}_j) f(\mathbf{x}_i) f(\mathbf{x}_j) \\
&\geq 0,
\end{aligned}$$

car K est d.p. $\square$

Preuve

On peut maintenant appliquer le *Théorème spectral* suivant à L_K:

Théorème 7 *Soit L un opérateur linéaire compact sur un espace de Hilbert $\mathcal{H}$. Alors il existe dans $\mathcal{H}$ un systeme orthonormal complet $(\psi_1, \psi_2, \ldots)$ de vecteurs propres de L. Les valeurs propres $(\lambda_1, \lambda_2, \ldots)$ sont réelles si L est auto-adjoint, et positives si L est positif.*

Dans le cas de L_K, les vecteurs propres ϕ_k associés aux valeurs propres $\lambda_k \neq 0$ sont des fonctions continues car:

$$\psi_k = \frac{1}{\lambda_k} L \psi_K$$

Preuve

On peut maintenant énoncer le *Théorème de Mercer*:

Théorème 8 *Soit $\mathcal{X}$ un espace normé compact, ν une mesure Borélienne sur $\mathcal{X}$ est K un noyau de Mercer. Soit $(\lambda_1, \lambda_2, \ldots)$ les valeurs propres décroissantes de L_K, et $(\psi_1, \psi_2, \ldots)$ les vecteurs propres correspondants. Alors, on a pour tous $\mathbf{x}, \mathbf{t} \in \mathcal{X}$:*

$$K(\mathbf{x}, \mathbf{t}) = \sum_{k=1}^{\infty} \lambda_k \psi_k(\mathbf{x}) \psi_k(\mathbf{t}),$$

ou la convergence est absolue pour chaque $\mathbf{x}, \mathbf{t} \in \mathcal{X}$, et uniforme sur $\mathcal{X} \times \mathcal{X}$.

Preuve

Du Théorème de Mercer on déduit le résultat suivant:

Théorème 9 *L'application*

$$
\begin{aligned}
\Phi : \mathcal{X} &\mapsto l^2 \\
\mathbf{x} &\mapsto \left(\sqrt{\lambda_k}\psi_k\left(\mathbf{x}\right)\right)_{k\in\mathbb{N}}
\end{aligned}
$$

est bien définie, continue, et satisfait:

$$
K\left(\mathbf{x},\mathbf{t}\right) = \left\langle \Phi\left(\mathbf{x}\right), \Phi\left(\mathbf{t}\right)\right\rangle_{l^2} .
$$

Preuve

Preuve du Thépreme 9:
Par le Théorème de Mercer on voit que pour tout $\mathbf{x} \in \mathcal{X}$, $\sum \lambda_k \psi_k^2(\mathbf{x})$ converge vers $K(\mathbf{x}, \mathbf{x}) < \infty$, et donc $\Phi(\mathbf{x}) \in l^2$.
La continuité de Φ découle de:

$$\begin{aligned} \| \Phi(\mathbf{x}) - \Phi(\mathbf{t}) \|_{l^2}^2 &= \sum_{k=1}^{\infty} \lambda_k \left(\psi_k(\mathbf{x}) - \psi_k(\mathbf{t}) \right)^2 \\ &= K(\mathbf{x}, \mathbf{x}) + K(\mathbf{t}, \mathbf{t}) - 2K(\mathbf{x}, \mathbf{t}) \end{aligned}$$

Bilan

Nous avons donc démontré qu'un noyau défini positif peut s'écrire comme un produit scalaire dans un espace de Hilbert si:

- $\mathcal{X}$ est fini, ou
- $\mathcal{X}$ est un espace métrique compact et K est continu.

Pour démontrer le résultat dans le cas général, nous introduisons une nouvelle construction explicite d'un espace de Hilbert, non basée sur la diagonalisation d'un opérateur: le r.k.h.s.

Espace de Hilbert à noyau reproduisant (rkhs)

Définitions

Soit $\mathcal{X}$ un espace quelconque, et $(\mathcal{H}, \langle ., . \rangle_{\mathcal{H}})$ un espace de Hilbert de fonctions ($\mathcal{H} \subset \mathbb{R}^{\mathcal{X}}$).

Définition 10 *Une fonction $K : \mathcal{X}^2 \mapsto \mathbb{R}$ est appelée un* noyau reproduisant (noté n.r.) *ssi:*

- *$\mathcal{H}$ contient toutes les fonctions de la forme*

$$\forall \mathbf{x} \in \mathcal{X}, \quad K_{\mathbf{x}} : \mathbf{t} \mapsto K(\mathbf{x}, \mathbf{t})$$

- *Pour tout $\mathbf{x} \in \mathcal{X}$ et $f \in \mathcal{H}$, on a:*

$$f(\mathbf{x}) = \langle f, K_{\mathbf{x}} \rangle_{\mathcal{H}}$$

Si un n.r. existe, $\mathcal{H}$ est appelé un espace de Hilbert à noyau reproduisant (rkhs).

Propriétés des n.r. et rkhs

Théorème 11 *(Aronszajn, 1950)*

- *Si un n.r. existe, il est unique.*
- *Un n.r. existe si et seulement si $\forall \mathbf{x} \in \mathcal{X}$, la fonctionnelle $f \mapsto f(\mathbf{x})$ (de $\mathcal{H}$ dans $\mathbb{R}$) est continue.*
- *Un n.r. est un noyau d.p.*
- *Réciproquement, si K est d.p., alors il existe un rkhs ayant K pour n.r.*
- *Si K est un n.r., il vérifie la propriété reproduisante:*

$$\forall (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^2, \quad \langle K_{\mathbf{x}}, K_{\mathbf{y}} \rangle_{\mathcal{H}} = K(\mathbf{x}, \mathbf{y}) .$$

Remarques

- Soit K un noyau d.p. sur un espace $\mathcal{X}$ quelconque. Par le point 4, on peut lui associer un rkhs $\mathcal{H}$. Soit $\Phi : \mathcal{X} \mapsto \mathcal{H}$ défini par $\Phi(\mathbf{x}) = K_{\mathbf{x}}$. Par le point 5, on a:

$$\forall (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^2, \quad K(\mathbf{x}, \mathbf{y}) = \langle \Phi(\mathbf{x}), \Phi(\mathbf{y}) \rangle_{\mathcal{H}}$$

  ce qui prouve le Théorème 2 dans le cas général.
- Le rkhs a de nombreuses autres propriétés que nous utiliserons par la suite pour développer des algorithmes puissants.

Preuve: unicité

Si K et K' sont deux n.r. sur $\mathcal{H}$, alors on a pour tout $\mathbf{x} \in \mathcal{X}$:

$$\begin{aligned}
\| K_{\mathbf{x}} - K'_{\mathbf{x}} \|^2_{\mathcal{H}} &= \left\langle K_{\mathbf{x}} - K'_{\mathbf{x}}, K_{\mathbf{x}} - K'_{\mathbf{x}} \right\rangle_{\mathcal{H}} \\
&= \left\langle K_{\mathbf{x}} - K'_{\mathbf{x}}, K_{\mathbf{x}} \right\rangle_{\mathcal{H}} - \left\langle K_{\mathbf{x}} - K'_{\mathbf{x}}, K'_{\mathbf{x}} \right\rangle_{\mathcal{H}} \\
&= K_{\mathbf{x}}\left(\mathbf{x}\right) - K'_{\mathbf{x}}\left(\mathbf{x}\right) - K_{\mathbf{x}}\left(\mathbf{x}\right) + K'_{\mathbf{x}}\left(\mathbf{x}\right) \\
&= 0
\end{aligned}$$

donc $K = K'$ (dans $\mathcal{H}$ et donc partout). $\square$

Preuve: continuité

Si un n.r. K existe, alors on a pour tout $(\mathbf{x}, f) \in \mathcal{X} \times \mathcal{H}$:

$$\begin{aligned} \mid f(\mathbf{x}) \mid &= \mid \langle f, K_{\mathbf{x}} \rangle_{\mathcal{H}} \mid \\ &\leq \parallel f \parallel_{\mathcal{H}} . \parallel K_{\mathbf{x}} \parallel_{\mathcal{H}} \text{ (Cauchy-Schwarz)} \\ &\leq \parallel f \parallel_{\mathcal{H}} . K(\mathbf{x}, \mathbf{x})^{\frac{1}{2}} , \end{aligned}$$

car $\parallel K_{\mathbf{x}} \parallel_{\mathcal{H}}^2 = \langle K_{\mathbf{x}}, K_{\mathbf{x}} \rangle_{\mathcal{H}} = K(\mathbf{x}, \mathbf{x})$. Donc $f \in \mathcal{H} \mapsto f(\mathbf{x}) \in \mathbb{R}$ est une application linéaire *continue*. $\square$

Preuve: continuité (réciproque)

Réciproquement, supposons que pour tout $\mathbf{x} \in \mathcal{X}$, la forme linéaire $f \in \mathcal{H} \mapsto f(\mathbf{x})$ soit continue.
Par le Théorème de représentation de Riesz, il existe un unique $g_{\mathbf{x}} \in \mathcal{H}$ tel que:

$$f(\mathbf{x}) = \langle f, g_{\mathbf{x}} \rangle_{\mathcal{H}}$$

La fonction $K(\mathbf{x}, \mathbf{y}) = g_{\mathbf{x}}(\mathbf{y})$ est alors un n.r. pour $\mathcal{H}$. $\square$

Preuve: un n.r. est d.p.

Un n.r. est *symétrique* car, pour tout $(\mathbf{x},\mathbf{y}) \in \mathcal{X}^2$:

$$K(\mathbf{x},\mathbf{y}) = \langle K_{\mathbf{x}}, K_{\mathbf{y}} \rangle_{\mathcal{H}} = \langle K_{\mathbf{y}}, K_{\mathbf{x}} \rangle_{\mathcal{H}} = K(\mathbf{y},\mathbf{x}) \, .$$

Il est *d.p.* car pour tout $N \in \mathbb{N}$,$(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \in \mathcal{X}^N$, et $(a_1, a_2, \ldots, a_N) \in \mathbb{R}^N$:

$$\begin{aligned}
\sum_{i,j=1}^{N} a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) &= \sum_{i,j=1}^{N} a_i a_j \left\langle K_{\mathbf{x}_i}, K_{\mathbf{x}_j} \right\rangle_{\mathcal{H}} \\
&= \| \sum_{i=1}^{N} a_i K_{\mathbf{x}_i} \|_{\mathcal{H}}^2 \\
&\geq 0. \quad \square
\end{aligned}$$

Preuve: un n.d.p. est un n.r.

Soit K un noyau d.p. Nous allons construire *explicitement* un espace de Hilbert de fonctions définies sur $\mathcal{X}$ qui admette K pour n.r.
Soit $\mathcal{H}_0$ le sous-espace vectoriel de $\mathbb{R}^{\mathcal{X}}$ engendré par les fonctions $\{K_{\mathbf{x}}\}_{\mathbf{x}\in\mathcal{X}}$. Pour $f, g \in \mathcal{H}_0$ s'écrivant:

$$f = \sum_{i=1}^{m} a_i K_{\mathbf{x}_i}, \quad g = \sum_{j=1}^{n} b_j K_{\mathbf{y}_i},$$

définissons:

$$\langle f, g \rangle_{\mathcal{H}_0} := \sum_{i,j} a_i b_j K\left(\mathbf{x}_i, \mathbf{y}_j\right).$$

Preuve: un n.d.p. est un n.r.(cont.)

Cette fonction *ne dépend pas* de la décomposition de f ou g, car:

$$\langle f, g \rangle_{\mathcal{H}_0} = \sum_{i=1}^{m} a_i g\left(\mathbf{x}_i\right) = \sum_{j=1}^{n} b_j f\left(\mathbf{y}_j\right).$$

Cela montre aussi que c'est une *forme bilinéaire symétrique*.
K étant d.p., on a d'autre part:

$$\| f \|_{\mathcal{H}_0}^2 = \sum_{i,j=1}^{m} a_i a_j K\left(\mathbf{x}_i, \mathbf{x}_j\right) \geq 0.$$

Preuve: un n.d.p. est un n.r.(cont.)

Cela montre aussi que pour tout $\mathbf{x} \in \mathcal{X}$ et $f \in \mathcal{H}_0$:

$$\langle f, K_{\mathbf{x}} \rangle_{\mathcal{H}_0} = f(\mathbf{x}) .$$

On en déduit la propriété de *reproductibilité*:

$$\forall (\mathbf{x}, \mathbf{y}) \in \mathcal{X}^2, \quad \langle K_{\mathbf{x}}, K_{\mathbf{y}} \rangle_{\mathcal{H}_0} = K(\mathbf{x}, \mathbf{y})$$

D'autre part, par Cauchy-Schwarz, on a $\forall \mathbf{x} \in \mathcal{X}$:

$$|\, f(\mathbf{x}) \,| = \left| \langle f, K_{\mathbf{x}} \rangle_{\mathcal{H}_0} \right| \leq \| \, f \, \|_{\mathcal{H}_0} . K(\mathbf{x}, \mathbf{x})^{\frac{1}{2}} ,$$

donc $\| \, f \, \|_{\mathcal{H}_0} = 0 \Leftrightarrow f = 0$.

Preuve: un n.d.p. et un n.r.(cont.)

$\langle .,. \rangle_{\mathcal{H}_0}$ est donc un *produit scalaire*, et $\left(\mathcal{H}_0, \langle .,. \rangle_{\mathcal{H}_0}\right)$ un *espace pré-Hilbertien*.

Soit $\mathcal{H}$ l'espace de Hilbert obtenu en *complétant* $\mathcal{H}_0$ par les limites des suites de Cauchy (dans $\mathcal{H}_0$). Si $(f_n)_{n \geq 0}$ est une telle suite de Cauchy, alors:

$$\forall \left(\mathbf{x}, m, n\right) \in \mathcal{X} \times \mathbb{N}^2, \quad \mid f_m\left(\mathbf{x}\right) - f_n\left(\mathbf{x}\right) \mid \leq \parallel f_m - g_n \parallel_{\mathcal{H}_0} . K\left(\mathbf{x}, \mathbf{x}\right)^{\frac{1}{2}} ,$$

donc en tout point $\mathbf{x}$ la suite $\left(f_n(\mathbf{x})\right)_{n \geq 0}$ est une suite de Cauchy dont on appelle la limite $f\left(\mathbf{x}\right)$ dans $\mathbb{R}$. En rajoutant ces fonctions f dans $\mathcal{H}_0$ et en étendant le produit scalaire par passage à la limite, on obtient un *espace de Hilbert* $\mathcal{H}$ de fonctions de $\mathcal{X}$ dans $\mathbb{R}$, qui admet K pour n.r. $\square$

Exemple: rkhs du noyau linéaire

Soit $\mathcal{X} = \mathbb{R}^d$ et $K(\mathbf{x}, \mathbf{y}) = \langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}^d}$ le noyau linéaire.
Le rkhs associé est constitué des fonctions de la forme:

$$\mathbf{x} \in \mathbb{R}^d \mapsto f(\mathbf{x}) = \sum_i a_i \langle \mathbf{x}_i, \mathbf{x} \rangle_{\mathbb{R}^d} = \langle \mathbf{w}, \mathbf{x} \rangle_{\mathbb{R}^d}$$

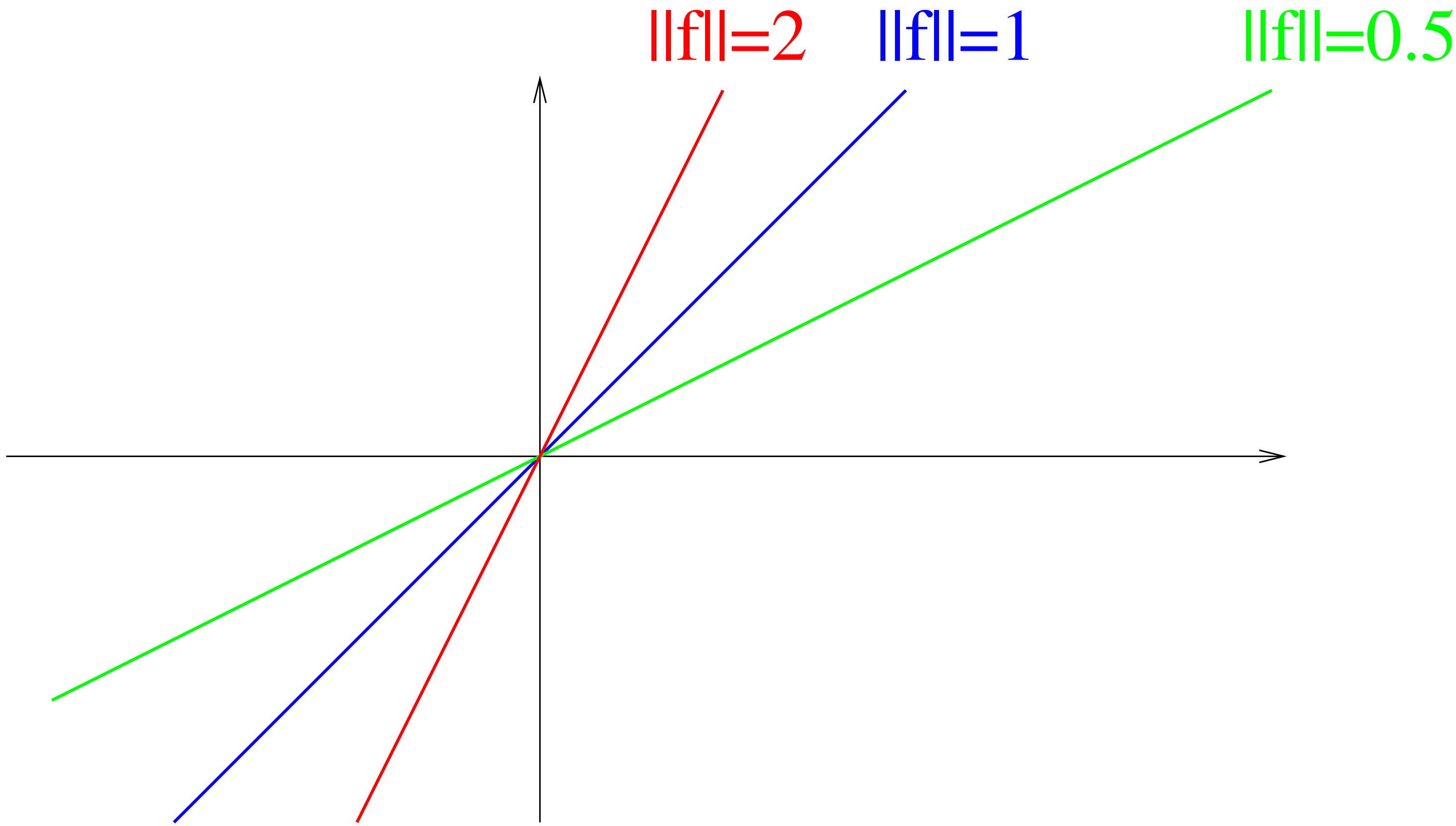
avec $\mathbf{w} = \sum_i a_i \mathbf{x}_i$. Le rkhs est donc l'ensemble des *formes linéaires*, muni du produit scalaire:

$$\langle f, g \rangle_{\mathcal{H}_K} = \langle \mathbf{w}, \mathbf{v} \rangle_{\mathbb{R}^d}$$

avec $f(\mathbf{x}) = \mathbf{w}.\mathbf{x}$ et $g(\mathbf{x}) = \mathbf{v}.\mathbf{x}$.

Exemple (cont.)

En particulier, $\| f \|_{\mathcal{H}_K} = \| \mathbf{w} \|_{\mathbb{R}^d}$, la pente de la forme linéraire.



Noyaux de Mercer et rkhs

Motivations

- Soit $\mathcal{X}$ un expace métrique compact, et K un noyau de Mercer (symétrique, d.p. et continu) sur $\mathcal{X}$.
- Dans cette partie nous faisons le lien entre le rkhs associé et la décomposition du noyau fourni par le Théorème de Mercer.

Rappel: Théorème de Mercer

Soit L_K l'opérateur linéaire sur $L_2^\nu(\mathcal{X})$ défini par:

$$\forall f \in L_2^\nu(\mathcal{X}), (L_K f)(\mathbf{x}) = \int K(\mathbf{x}, \mathbf{t}) f(\mathbf{t}) d\nu(\mathbf{t}).$$

Soient $(\lambda_1, \lambda_2, \ldots)$ les valeurs propres décroissantes de L_K, et $(\psi_1, \psi_2, \ldots)$ les vecteurs propres correspondants. Alors on a pour tous $\mathbf{x}, \mathbf{y} \in \mathcal{X}$:

$$K(\mathbf{x}, \mathbf{y}) = \sum_{k=1}^{\infty} \lambda_k \psi_k(\mathbf{x}) \psi_k(\mathbf{y}) = \langle \Phi(\mathbf{x}), \Phi(\mathbf{y}) \rangle_{l^2},$$

avec $\Phi : \mathcal{X} \mapsto l^2$ défini par $\Phi(\mathbf{x}) = \left(\sqrt{\lambda_k} \psi_k(\mathbf{x})\right)_{k \in \mathbb{N}}$.

Construction du rkhs

Nous supposons que $\lambda_k > 0$ pour tout $k \geq 1$ (sinon, le résultat et la preuve restent valide dans le sous-espace engendré par les vecteurs propres de valeur propres non nulles).
Soit l'espace de Hilbert:

$$H_K = \left\{ f \in L_2^{\nu}(\mathcal{X}) : f = \sum_{i=1}^{\infty} a_i \psi_i, \quad \text{avec } \sum_{k=1}^{\infty} \frac{a_k^2}{\lambda_k} < \infty \right\}$$

muni du produit scalaire:

$$\langle f, g \rangle_K = \sum_{k=1}^{\infty} \frac{a_k b_k}{\lambda_k}, \quad \text{pour } f = \sum_k a_k \psi_k, g = \sum_k b_k \psi_k.$$

Construction du rkhs (cont.)

Pour montrer que H_K est le rkhs associé au noyau K, nous devons montrer que:

- c'est un *espace de fonctions* de $\mathcal{X}$ dans $\mathbb{R}$,
- pour chaque $\mathbf{x} \in \mathcal{X}$, $K_x \in H_K$,
- pour chaque $\mathbf{x} \in \mathcal{X}$ et $f \in H_K$, $f(\mathbf{x}) = \langle f, K_x \rangle_{H_K}$.

Construction du rkhs (cont.)

Si $f = \sum_{i=1}^{\infty} a_i \psi_i \in H_K$ et $\mathbf{x} \in \mathcal{X}$, on a:

$$
\begin{aligned}
\mid f(\mathbf{x}) \mid &= \left| \sum_{i=1}^{\infty} a_i \psi_i(\mathbf{x}) \right| = \left| \sum_{i=1}^{\infty} \frac{a_i}{\sqrt{\lambda_i}} \sqrt{\lambda_i} \psi_i(\mathbf{x}) \right| \\
&\leq \left(\sum_{i=1}^{\infty} \frac{a_i^2}{\lambda_i} \right)^{\frac{1}{2}} \cdot \left(\sum_{i=1}^{\infty} \lambda_i \psi_i(\mathbf{x})^2 \right)^{\frac{1}{2}} \\
&= \| f \|_{H_K} . K(\mathbf{x}, \mathbf{x})^{\frac{1}{2}} .
\end{aligned}
$$

donc la convergence dans $\| . \|_{H_K}$ implique la convergence ponctuelle.

Construction du rkhs (cont.)

On en déduit que toute fonction $f = \sum_{i=1}^{\infty} a_i \psi_i \in H_K$ est bien une fonction définie ponctuellement, et continue (comme les ψ_i).
Soit maintenant $\mathbf{x} \in \mathcal{X}$. Pour montrer que $K_x \in H_K$, soit $a_i = \lambda_i \psi_i\left(\mathbf{x}\right)$. On a bien:

$$\sum_{i=1}^{\infty} \frac{a_i^2}{\lambda_i} = \sum_{i=1}^{\infty} \lambda_i \psi_i\left(\mathbf{x}\right)^2 = K\left(\mathbf{x}, \mathbf{x}\right) < \infty,$$

donc $\phi_x := \sum_{i=1}^{\infty} a_i \psi_i \in H_K$.

Construction du rkhs (cont.)

Par le précédent résultat, cette convergence dans H_K a aussi lieu ponctuellement, donc pour tout $\mathbf{t} \in \mathcal{X}$:

$$\phi_x(\mathbf{t}) = \sum_{i=1}^{\infty} a_i \psi_i(\mathbf{t}) = \sum_{i=1}^{\infty} \lambda_i \psi_i(\mathbf{x}) \psi_i(\mathbf{t}) = K(\mathbf{x}, \mathbf{t}),$$

donc $\phi_x = K_x \in H_K$.
Finalement, pour $g = \sum_{i=1}^{\infty} b_i \psi_i \in H_K$, on a:

$$\langle g, K_x \rangle_{H_K} = \langle g, \phi_x \rangle_{H_K} = \sum_{i=1}^{\infty} \frac{\lambda_i \psi_i(\mathbf{x}) b_i}{\lambda_i} = f(\mathbf{x}),$$

ce qui termine la preuve. $\square$

Remarques

- Bien que H_K ait été construit à partir des fonctions propres de L_K, lui-meme défini à partir de la mesure $\nu(\mathbf{x})$, ce résultat montre que *H_K est indépendant de ν et L_K*.
- Le Théorème de Mercer fournit une *manière concrète* de construire un rkhs, en prenant des combinaisons linéaires des fonctions propres de L_K (et la condition sur les coefficients).
- Les $(\psi_i)_{i\in\mathbb{N}}$ forment une *base orthogonale* du rkhs:

$$\langle \psi_i, \psi_j \rangle_{H_K} = 0 \quad \text{si } i \neq j, \quad \| \psi_i \|_{H_K} = \frac{1}{\sqrt{\lambda_i}}.$$

Rkhs et régularisation par fonctions de Green

Motivation

- A tout noyau d.p. est associé un espace de Hilbert fonctionnel: le rkhs
- Dans le cas $\mathcal{X}$ compact, le Théorème de Mercer fournit une premiere intuition sur cet espace.
- Quid de cas plus généraux, par exemple $\mathcal{X} = \mathbb{R}^d$?
- Nous allons montrer que dans certains cas, on peut interpréter $\| f \|_{\mathcal{H}}$ comme une *mesure de la régularité* de f.

Exemple

Soit

$\mathcal{H} = \{ f : [0,1] \mapsto \mathbb{R},$ cont., dérivable p.p., $f' \in L^2\left([0,1]\right), f(0) = 0$

C'est un *espace de Hilbert* quand on le muni du produit scalaire:

$$\forall (f,g) \in \mathcal{F}^2 \ \langle f, g \rangle_{\mathcal{H}} = \int_0^1 f'(u)\, g'(u)\, du.$$

La norme $\| f \|_{\mathcal{H}}$ mesure la régularité de f:

$$\| f \|_{\mathcal{H}} = \left(\int_0^1 f'(u)^2 du \right)^{\frac{1}{2}} = \| f' \|_{L^2([0,1])}.$$

Exemple (cont.)

Théorème 12 $\mathcal{H}$ est un rkhs *ayant pour n.r. la fonction:*

$$\forall (x,y) \in \mathcal{X}^2, \quad K(x,y) = \min(x,y).$$

Comme $\| f \|_{\mathcal{H}} = \| f' \|_{L^2}$, cela montre que la norme dans le rkhs est justement cette fonctionnelle de régularité.

Exemple (cont.)

Pour prouver ce Théorème, nous allons montrer que:

- Les fonctionnelles d'évaluation sont continues,
- $\forall x \in [0,1]$, $K_x \in \mathcal{F}$,
- $\forall (x,f) \in [0,1] \times \mathcal{H}, \langle f, K_x \rangle_{\mathcal{H}} = f(x)$.

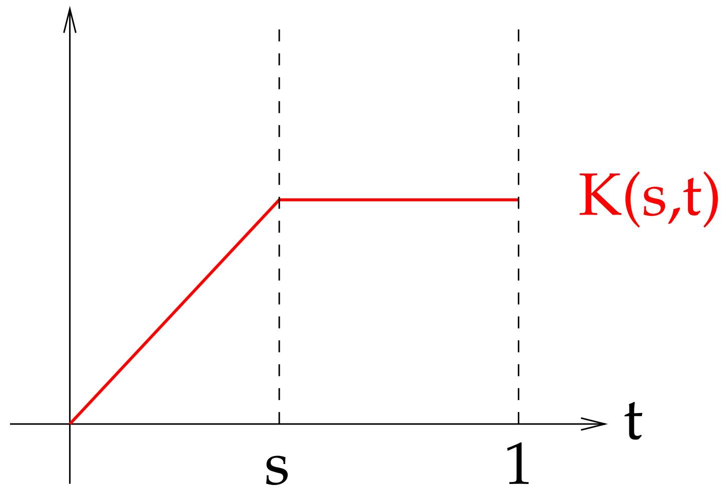
La premiere propriété découle de:

$$f(x) = f(x) - f(0) = \int_0^t f'(u)du \leq \sqrt{t} \| f \|_{\mathcal{H}},$$

valable pour tout $x \in [0,1]$ et $f \in \mathcal{H}$.

Exemple (cont.)

Pour la seconde propriété, soit $K_x(y) = K(x,y) = \min(x,y)$ sur $[0,1]^2$.



K_x est dérivable sauf en s, a une dérivée de carré intégrable, et $K_x(0) = 0$, donc $K_x \in \mathcal{H}$ pour tout $x \in [0,1]$.

Exemple (cont.)

Enfin, pour la troisieme propriété, observons que l'égalité suivante est valable pour tout $x \in [0,1]$ et $f \in \mathcal{H}$:

$$\langle f, K_x \rangle_{\mathcal{H}} = \int_0^1 f'(u) K_x'(u) du = \int_0^x f'(u) du = f(x),$$

ce qui montre que K est le n.r. associé à $\mathcal{H}$. $\square$

Généralisation

Soit $\mathcal{X} = \mathbb{R}^d$ et D un opérateur différentiel sur une classe de fonctions $\mathcal{H}$ telle que munie du produit scalaire:

$$\forall (f,g) \in \mathcal{H}^2, \quad \langle f,g \rangle_{\mathcal{H}} = \langle Df, Dg \rangle_{L^2(\mathcal{X})},$$

ce soit un espace de Hilbert.

Théorème 13 *$\mathcal{H}$ est un rkhs dont le n.r. est la fonction de Green de l'opérateur D^*D, ou D^* est l'adjoint de D.*

Rappel: fonction de Green

Soit l'équation différentielle sur $\mathcal{H}$:

$$f = Dg,$$

ou g est l'inconnu. Pour la résoudre, on peut chercher g de la forme:

$$g(x) = \int_{\mathcal{X}} k(x,y) f(y) \, dy$$

pour une certaine fonction $k : \mathcal{X}^2 \mapsto \mathbb{R}$. k doit vérifier, pour tout $x \in \mathcal{X}$,

$$f(x) = Dg(x) = \langle Dk_x, f \rangle_{L^2(\mathcal{X})}$$

k est appelée *fonction de Green* de l'opérateur D.

Preuve du Théorème 13

Soit donc $\mathcal{H}$ un espace de Hilbert muni du produit scalaire:

$$\langle f, g \rangle_{\mathcal{X}} = \langle Df, Dg \rangle_{L^2(\mathcal{X})} ,$$

et K la fonction de Green de l'opérateur D^*D.
Pour tout $x \in \mathcal{X}$, $K_x \in \mathcal{H}$ car:

$$\langle DK_x, DK_x \rangle_{L^2(\mathcal{X})} = \langle D^*DK_x, K_x \rangle_{L^2(\mathcal{X})} = K_x(x) < \infty .$$

D'autre part, pour tout $f \in \mathcal{H}$ et $x \in \mathcal{X}$, on a:

$$f(x) = \langle D^*DK_x, f \rangle_{L^2(\mathcal{X})} = \langle DK_x, Df \rangle_{L^2(\mathcal{X})} = \langle K_x, f \rangle_{\mathcal{H}}$$

ce qui montre $\mathcal{H}$ est un rkhs dont le n.r. est K. $\square$

Noyaux et régularisation par transformée de Fourier

Mercer sur domaines non borné

Supposons $\mathcal{X}$ non compact, par exemple $\mathcal{X} = \mathbb{R}^d$. Dans ce cas, les valeurs propres de l'équation:

$$\int_{\mathcal{X}} K\left(\mathbf{x},\mathbf{t}\right)\psi\left(\mathbf{t}\right) = \lambda\psi\left(\mathbf{t}\right)$$

ne sont pas nécessairement dénombrables, le Théorème de Mercer n'est pas valable.

Rappel: transformée de Fourier

Définition 14 *Soit $f \in L^1\left(\mathbb{R}^d\right)$.*

- *On appelle* transformée de Fourier *de f, et note $\hat{f}$ ou $\mathcal{F}f$, la fonction définie pour tout $\omega \in \mathbb{R}^d$ par:*

$$\hat{f}\left(\omega\right) = \int_{\mathbb{R}^d} e^{-ix.\omega} f\left(x\right) dx.$$

- *$\hat{f}$ est à valeur complexe, continue, tend vers 0 à l'infini et $\| \hat{f} \|_{L^\infty} \leq \| f \|_{L^1}$.*
- *Si en plus $\hat{f} \in L^1\left(\mathbb{R}^d\right)$, alors on a la formule d'*inversion de Fourier*:*

$$\forall x \in \mathbb{R}^d, \quad f\left(x\right) = \frac{1}{\left(2\pi\right)^d} \int_{\mathbb{R}^d} e^{ix.\omega} \hat{f}\left(\omega\right) d\omega.$$

Rappel: formule de Parseval

Théorème 15 *Si $f \in L^1\left(\mathbb{R}^d\right)$ est en plus de carré sommable, alors la* formule de Parseval *est valide:*

$$\int_{\mathbb{R}^d} |\, f(x)\,|^2 \, dx = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \left|\, \hat{f}(\omega) \right|^2 d\omega$$

Noyaux invariants par translation (i.t.)

Définition 16 *Un noyau $K : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$ est dit* invariant par translation *s'il ne dépend que de la différence entre ses arguments, i.e.:*

$$\forall (x, y) \in \mathbb{R}^{2d}, \quad K(x, y) = K(x - y).$$

La plupart des noyaux sur $\mathbb{R}^d$ que nous rencontrerons sont i.t.

Mercer pour noyaux i.t.?

Soit K un noyau i.t. sommable. On a alors, pour tout $(x, y) \in \mathbb{R}^{2d}$:

$$K(x-y) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i(x-y).\omega} \hat{K}(\omega)\, d\omega$$
$$= \int_{\mathbb{R}^d} \frac{\hat{K}(\omega)}{(2\pi)^d} e^{i\omega(x)} e^{i\omega(-y)} d\omega.$$

Cela ressemble à la décomposition d'un noyau de Mercer sur un domaine compact:

$$K(\mathbf{x}, \mathbf{y}) = \sum_i \lambda_i \psi_i(\mathbf{x}) \psi_i(\mathbf{y})$$

Rkhs pour noyaux i.t.

Théorème 17 *Soit K un noyau d.p., i.t., sommable et de transformée de Fourier $\hat{K}$ sommable sur $\mathbb{R}^d$. Le sous-espace $\mathcal{H}_K$ de $L_2\left(\mathbb{R}^d\right)$ des fonctions f continues telles que:*

$$\| f \|_K^2 := \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \frac{\left| \hat{f}(\omega) \right|^2}{\hat{K}(\omega)} d\omega < +\infty,$$

muni du produit scalaire:

$$\langle f, g \rangle := \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \frac{\hat{f}(\omega)\hat{g}(\omega)^*}{\hat{K}(\omega)} d\omega$$

est un rkhs avec K pour n.r.

Preuve du Théorème 17

Pour $x \in \mathbb{R}^d$, $K_x(y) = K(x - y)$ donc:

$$\hat{K_x}(\omega) = \int e^{-i\omega.u} K(u - x) du = e^{-i\omega.x} \hat{K}(\omega).$$

On en déduit que $K_x \in \mathcal{H}$, car:

$$\int_{\mathbb{R}^d} \frac{\left| \hat{K_x}(\omega) \right|^2}{\hat{K}(\omega)} \leq \int_{\mathbb{R}^d} \left| \hat{K}(\omega) \right| < \infty,$$

De plus, si $f \in \mathcal{H}$ et $x \in \mathbb{R}^d$, on a:

$$\langle f, K_x \rangle_{\mathcal{H}} = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \frac{\hat{K_x}(\omega) \hat{f}(\omega)^*}{\hat{K}(\omega)} d\omega = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \hat{f}(\omega)^* e^{-i\omega.x} = f($$

Application: Théorème de Bochner

Théorème 18 *Une fonction* $K(x-y)$ *sur* $\mathbb{R}^d$ *est définie positive si et seulement si elle est la transformée de Fourier d'une fonction* $\hat{K}(\omega)$ symétrique, positive, et tendant vers 0 à l'infini*.*

Exemple: noyau Gaussien

$$K(x,y) = e^{-\frac{(x-y)^2}{2\sigma^2}}$$

correspond a:

$$\hat{K}(\omega) = e^{-\frac{\sigma^2\omega^2}{2}}$$

et

$$\mathcal{H} = \left\{ f : \int \left| \hat{f}(\omega) \right|^2 e^{\frac{\sigma^2\omega^2}{2}} d\omega < \infty \right\},$$

l'ensemble des fonctions infiniment différentiables dont toutes les dérivées sont L^2.

Exemple: noyau de Laplace

$$K(x,y) = \frac{1}{2} e^{-\gamma | x-y |}$$

correspond a:

$$\hat{K}(\omega) = \frac{\gamma}{\gamma^2 + \omega^2}$$

et

$$\mathcal{H} = \left\{ f : \int \left| \hat{f}(\omega) \right|^2 \frac{\left(\gamma^2 + \omega^2\right)}{\gamma} d\omega < \infty \right\},$$

l'ensemble des fonctions L^2 differentiables dont la dérivée est L^2 (Sobolev).

Exemple: noyau passe-bande

$$K(x,y) = \frac{\sin(\Omega(x-y))}{\pi(x-y)}$$

correspond a:

$$\hat{K}(\omega) = U(\omega+\Omega) - U(\omega-\Omega)$$

et

$$\mathcal{H} = \left\{ f : \int_{|\omega|>\Omega} \left| \hat{f}(\omega) \right|^2 d\omega = 0 \right\},$$

l'ensemble des fonctions dont le spectre est inclus dans $[-\Omega, \Omega]$.