### 2.1.2 Cadre fonctionnel

Un algorithme d'apprentissage permet de choisir, dans un ensemble d'hypothèses $\mathscr{H}$, la meilleure solution compte tenu de la base d'apprentissage et des deux critères à optimiser. Les méthodes à base de noyaux construisent cet ensemble d'hypothèses $\mathscr{H}$ à l'aide d'un noyau. Intuitivement, pour classer des points, il est utile de disposer d'une notion de proximité ou de ressemblance entre les objets que l'on cherche à séparer. Si l'on sait donner une distance entre deux objets, quelque soit leur forme, on peut se contenter de cette information pour réaliser la discrimination. En partant de cette idée, nous définissons les *noyaux* :

**DÉFINITION** - UN NOYAU : est une fonction symétrique de deux variables qui retourne un scalaire exprimant une notion de distance entre les variables. Plus formellement, soient $\mathbf{u},\mathbf{v} \in \mathbb{R}^d$, on défini $k(.,.)$ comme étant :

$$\begin{array}{cccc} k(.,.) & : & \mathbb{R}^d \times \mathbb{R}^d & \rightarrow & \mathbb{R} \\ & & \mathbf{u},\mathbf{v} & \mapsto & k(\mathbf{u},\mathbf{v}) \end{array}$$

A partir du noyau (lorsqu'il est défini positif[2]), on peut construire l'espace des hypothèses en passant par l'espace pré-hilbertien $\mathscr{H}_0$.

**DÉFINITION** - ESPACE PRÉ-HILBERTIEN. : Soit $k(,)$ un noyau défini positif. On définit $\mathscr{H}_0$ l'espace vectoriel engendré par les combinaisons linéaires finies du noyau :

$$\mathscr{H}_0 = \left\{ f : \mathbb{R}^d \mapsto \mathbb{R} \;\middle|\; \ell < \infty, \{\alpha_i\}_1^\ell \in \mathbb{R}, \{\mathbf{u}_i\}_1^\ell \in \mathbb{R}^d, f(\mathbf{x}) = \sum_{i=1}^{\ell} \alpha_i k(\mathbf{u}_i,\mathbf{x}) \right\}$$

On donne à cet espace le produit scalaire défini par la forme bilinéaire telle que pour tout couple $f,g \in \mathscr{H}_0, f(\mathbf{x}) = \sum_{i=1}^{\ell} \alpha_i k(\mathbf{u}_i,\mathbf{x})$ et $g(\mathbf{x}) = \sum_{j=1}^{m} \beta_j k(\mathbf{v}_j,\mathbf{x})$,

$$\langle f,g \rangle_{\mathscr{H}_0} = \sum_{i=1}^{\ell} \sum_{j=1}^{m} \alpha_i \beta_j k(\mathbf{u}_i,\mathbf{v}_j)$$

La positivité du noyau garantit la positivité du produit scalaire ainsi défini. L'espace engendré muni de ce produit scalaire est un pré-hilbertien. La complétion de l'espace pré-hilbertien $\mathscr{H}_0$ au sens de la norme induite par le produit scalaire, est un espace de Hilbert noté $\mathscr{H} = \overline{\mathscr{H}_0}$.

**PROPRIÉTÉ :** La norme induite dans cet espace est $\|f\|^2_{\mathscr{H}} = \langle f,f \rangle_{\mathscr{H}}$. Par ailleurs, on remarque que :

$$\langle f(.),k(.,\mathbf{x}) \rangle_{\mathscr{H}} = \left\langle \sum_{i=1}^{\ell} \alpha_i k(\mathbf{u}_i,.),k(.,\mathbf{x}) \right\rangle_{\mathscr{H}} = f(\mathbf{x})$$

C'est la propriété de reproduction. L'espace $\mathscr{H}$ est donc un espace de Hilbert à noyaux reproduisant, aussi appelé *RKHS* (Reproducing Kernel Hilbert Space).

Grâce à l'utilisation judicieuse de la propriété de reproduction, le problème de recherche d'une solution générale $f$ dans l'espace de fonctions $\mathscr{H}$ se ramène à la recherche d'un vecteur $\alpha \in \mathbb{R}^n$ où $n$ désigne le nombre d'exemples disponibles pour l'apprentissage. Le lien entre la fonction $f$ et le vecteur $\alpha$ est donné par un théorème dit de représentation (théorème 4.2 page 90 [Schölkopf and Smola, 2002]) :

$$f(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i k(\mathbf{x}_i,\mathbf{x})$$

[2]Nous ne considérerons que le cas des noyaux définis positifs *i.e.* tels que $\forall \ell < \infty, \forall \alpha_i \in \mathbb{R}$, $\forall \mathbf{x}_i \in \mathbb{R}^d, i = 1...\ell$; $\sum_i \sum_j \alpha_i \alpha_j k(\mathbf{x}_i,\mathbf{x}_j) \geq 0$.