# MVA "Kernel methods"
# Homework 2

Jean-Philippe Vert

Due February 6, 2013

**Exercice 1.**
Given two sets of real numbers $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ and $Y = (y_1, \ldots, y_n) \in \mathbb{R}^n$, the covariance between $X$ and $Y$ is defined as

$$cov_n(X,Y) = \mathbf{E}_n(XY) - \mathbf{E}_n(X)\mathbf{E}_n(Y)\,,$$

where $\mathbf{E}_n(U) = (\sum_{i=1}^n u_i)/n$. The covariance is useful to detect linear relationships between $X$ and $Y$. In order to extend this measure to potential nonlinear relationships between $X$ and $Y$, we consider the following criterion:

$$C_n^K(X,Y) = \max_{f,g\in\mathcal{B}_K} cov_n(f(X), g(Y))\,,$$

where $K$ is a positive definite kernel on $\mathbb{R}$, $\mathcal{B}_K$ is the unit ball of the RKHS of $K$, and $f(U) = (f(u_1), \ldots, f(u_n))$ for a vector $U = (u_1, \ldots, u_n)$.
**1.** Express simply $C_n^K(X,Y)$ for the linear kernel $K(a,b) = ab$.
**2.** For a general kernel $K$, express $C_n^K(X,Y)$ in terms of the Gram matrices of $X$ and $Y$.